DOCUMENT RESUME

ED 277 164 EC 190 996

TITLE	The Miller Assessment for Preschoolers: A Longitudinal and Predictive Study. Final Report.
INSTITUTION	Foundation for Knowledge in Development, Littleton, CO.
SPONS AGENCY	American Occupational Therapy Association, Rockville, Md.; Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE	Jun 86
GRANT	1232116970A1
NOTE	113p.; Some tables contain small print.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Behavior Problems; Discriminant Analysis; *Handicap Identification; *High Risk Students; *Learning Problems; Longitudinal Studies; Predictive Measurement; *Predictive Validity; Predictor Variables; Preschool Education; Preschool Tests; School Readiness Tests; Standardized Tests; *Test Validity
IDENTIFIERS	*Miller Assessment for Preschoolers

ABSTRACT

The study reported here sought to establish the predictive validity of the Miller Assessment for Preschoolers (MAP), an instrument designed to identify preschool children at risk for school-related problems in the primary years. Children (N=338) in 11 states who were originally tested in 1980 as part of the MAP standardization project were given a 4-hour battery of standardized tests. Standardized criterion measures included the Wechsler Intelligence Scale-Revised, the Berry Developmental Test of Visual Motor Integration, the Bruininks-Oseretsky Test of Motor Proficiency, the Goodenough-Harris Drawing Test, the Woodcock-Johnson Psychoeducational Battery, and the Walker Problem Behavior Identification Checklist. Parents and teachers completed behavior checklists and history forms. Among subjective criterion measures were retention in school; teacher observations; receipt of special services; and report card grades in language, reading, mathematics, and physical education. Predictor variables examined were the MAP Total Score, the five MAP subtest scores, and the 27 MAP items. Data were subjected to correlational, T-test, stepwise multiple regression, and classificational analyses. Detailed results indicate that the predictive validity of the MAP compares favorably to similar predictive validity studies, and that the instrument can consistently and correctly identify the majority of children who later have difficulty in school, with relatively few misclassifications. (JW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277164

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

THE MILLER ASSESSMENT FOR PRESCHOOLERS:
A LONGITUDINAL AND PREDICTIVE STUDY

A FINAL REPORT TO:

THE U.S. DEPARTMENT OF EDUCATION
OFFICE OF SPECIAL EDUCATION AND REHABILITATIVE SERVICES
HANDICAPPED RESEARCH AND DEMONSTRATION (84.023)
1232116970A1
JUNE 1986

Lucy Jane Miller, Ph. D.

Foundation for Knowledge Development
4857 S. Albion Street
Littleton, Colorado 80121

EC190996

BEST COPY AVAILABLE

Acknowledgments

Numerous people contributed to this study since it's beginning in 1972, most of whom will go unacclaimed here. However, special thanks is given to the Field Staff of this project; Joy Morse, M.S., Evelyn Jaffe, M.S, Harriet Schmid, Ph.D, Sara Hurwitz, M.S., Richard Trifiletti, M.S., Eleanor Donahue, M.S., Linda McClain, M.S., Teresa Norris, M.S., Susan Smith, M.S., Marilyn Gora, M.S., Ben Atchison, M.S., Joseph Mastropaolo, M.S., Judy Su, M.S., OTR, to Tracy Sprong, Project Administrator; to Barb Bieber for consultation; and to Jim Zumbrennon for data analysis.

In addition, thanks is given to Toni Linder, Ray Kluever Bill Slaichert, and Chip Reichardt for academic guidance with this project.

As always with works of this sort, the previous efforts of several authors has made a significant contribution to the dissertation. In this dissertation these particular contributions were made by: Pamela Lemerand, Samual Meisels and Robert Lichtenstein.

Particular thanks is noted to Danny Miller, without whose computer expertise this would not have been finished.

Finally, I am, as always, thankful to Bill and Nicole whose encouragement and support sustain these efforts.

This project was funded by the U.S. Department of Education Office of Special Education and Rehabilitative Services (Grant #1232116970A1) and by a grant from the American Occupational Therapy Foundation.

FINAL REPORT
MILLER ASSESSMENT FOR PRESCHOOLERS
LONGITUDINAL STUDY

This final report provides a summary of this research study, a longitudinal predictive study of the Miller Assessment for Preschoolers (MAP). The report is divided into the following sections: A) Purpose of this Study; B) Background to this Project; C) Predictor and Criterion Variables; D) Methodology and Procedures; E) Research Questions and Conclusions; F) Limitations of this Study; G) Recommendations for Further Study; and H) Conclusion.

A. Purpose of this Study

The focus of this research was to establish the predictive validity of the Miller Assessment for Preschoolers (MAP). The need for quantifiable psychometric data on tests is well established in the literature. The MAP demonstrates many excellent qualities, such as extensive item development, careful standardization, significant reliability studies, and construct and content validity. However, prior to this study, the predictive validity of the MAP was unproven. Thus the potential usefulness of the MAP for preschool screening programs could not be evaluated.

B. Background to this Project

The development of the MAP began in 1972, with the examination of 115 preschool tests, review of 177 sources of research and theory, and preparation of a pilot edition of the MAP. From 1974 - 1977 three separate editions of the MAP were field tested, and additional item research was undertaken. In 1979 a Research Edition of the MAP was tested nationwide on a stratified, randomly selected sample (n = 600). In 1980, the MAP was standardized nationwide utilizing a randomly selected, stratified sample (n=1204). At this time reliability and content and construct validity studies were also implemented. Initial data on the MAP was promising; however in the absence of predictive validity data the value of the MAP in fulfilling its stated purpose (i.e. the identification in the preschool years of children at risk for school related problems in the primary years), could not be determined.

C. Predictor and Criterion Variables

The predictor variables in this study were all related to the MAP. Examined were: the MAP Total Score; the five MAP subtest scores: Foundations Index, Coordination Index, Verbal Index, Non-verbal Index, and Complex Tasks Index; and the twenty seven MAP items: Articulation, Block Designs, Block Tapping, Cage, Digit Repetition, Draw-A-Person, Figure Ground, Finger Localization, Follow Directions, General Information, Hand-Nose, Imitation of Postures, Kneel-Stand, Maze, Object Memory, Puzzles, Romberg, Sentence Repetition, Sequencing, Stamp, Stepping, Stereognosis, Supine Flexion, Tongue Movements, Tower, Vertical Writing, and Walks Line.

The criterion variables included both standardized assessments and more subjective measures. The standardized criterion measures were:

The Weschler Intelligence Scale-Revised
The Berry Developmental Test of Visual Motor Integration
The Bruininks-Oseretsky Test of Motor Proficiency
The Goodenough-Harris Drawing Test
The Woodcock-Johnson Psychoeducational Battery
The Walker Problem Behavior Identification Checklist

In addition the following more subjective criterion measures were employed to answer some of the research questions:

Retention in School
Teachers' Observations
Receipt of Special Services
Need for a Special Class
Report Card Grades in Language, Reading, Math and Physical Education
The Harter Scale of Perceived Competance and Acceptance

D. Methodology and Procedures

From 1980-1984 the children who were originally tested as part of the MAP standardization project were "tracked"; that is, the name and address list was kept as current as possible.

In the summer of 1984, 13 qualified professionals who were school psychologists or occupational therapists were hired to be Field Staff for this project. They were trained in a ten day seminar in assessment techniques and data gathering procedures. After reliability of administration and scoring were acceptable, Field Staff returned to the 11 states they represented, and began to locate and test children. Each child received a four hour battery of standardized tests. The tests were administered in the same order to each child, but the Field Staff rotated the test with which they began. In addition, parents and teachers filled out standardized behavior checklists and history forms. The Field Staff also gathered as much history as possible from the child's school records.

All data was sent to the main office in Denver where it was carefully checked for accuracy. Information was then translated via a code book onto coding sheets for data entry. Data was keypunched and checked for accuracy. Data analyses were then completed.

E. Research Questions and Conclusions

This section of the report presents the results of this study of the predictive validity of the Miller Assessment for Preschoolers with a sample of 338 children. Before addressing the research questions descriptive statistics concerning the outcome of the sample on the screening and the criterion measures are reported. The remainder of this section will consist of a statement of each research question, a brief description of the analyses used, presentation of the findings for each question, and a discussion of the results. The subsections are organized as follows:

Descriptive Statistics

Research Question #1: Correlational Analyses

1) MAP Total Score
2) MAP Subtest Scores
3) MAP Item Scores

Research Question #2: T-tests for Differences Between Means of Each of the Problem Categories

1) Retained in School
2) Failed Teachers' Observations
3) Received Special Services
4) In a Special Class
5) Failed Report Card Language
6) Failed Report Card Reading
7) Failed Report Card Math
8) Failed Report Card Physical Education

Research Question #3: Step-wise Multiple Regression for Each of the Problem Categories

1) Retained in School
2) Failed Teachers' Observations
3) Received Special Services
4) In a Special Class
5) Failed Report Card Language

6) Failed Report Card Reading
7) Failed Report Card Math
8) Failed Report Card Physical Education

Research Question #4: Summary of Analyses for # 1 - 3 with Focus on MAP Items

Research Question #5: Classificational Analysis for Eight Problem Categories and Four Standardized Measures

1) Retained in School
2) Failed Teachers' Observations
3) Received Special Services
4) In a Special Class
5) Failed Report Card Language
6) Failed Report Card Reading
7) Failed Report Card Math
8) Failed Report Card Physical Education
9) WISC-R
10) Woodcock Language
11) Woodcock Reading
12) Woodcock Math

Descriptive Statistics

Distribution of the sample by final percentile score on the MAP is presented in Table 1. It can be seen that the sample for this study was fairly normally distributed: 4.4 percent of the sample received final scores on the MAP between the 0 and 5th percentile; 23.9 percent of the sample received a final score on the MAP between the 0 and 25th percentile (first quartile); 52.3 percent of the sample received final scores on the MAP between the 0 and the 50th percentile (second quartile); 88.4 percent of the sample received a final score on the MAP between the 0 and the 75th percentile (third quartile); and the remainder of the sample received final scores on the MAP between the 0 and the 99th percentile (fourth quartile).

The distribution of scores on the standardized criterion measures is noted in Table 2. Of note is that all the standardized tests demonstrated means and standard deviations similar to the norms represented in the test standardization samples except the WISC-R and the Bruininks.

The WISC-R scores are approximately one standard deviation above the mean of the scores of the norm sample;

thus a score of 115 was average for this sample. The mean score of this sample on the Bruininks was just slightly less than one standard deviation above the mean of the scores of the norming sample. The standard deviations of the sample on all measures in this study were similar to the standard deviations of the norming samples.

It is not known why the WISC-R scores were skewed. One possible hypothesis for the cause of the high WISC-R scores, suggested by several psychologists who worked on this project is that the WISC-R no longer has a mean score of 100. In other words, the average performance of U.S. children in 1986 is closer to a score of 115 than 100. (It would be interesting to subject this hypothesis to study.)

The rest of this section presents the findings and discussion of results for each of the research questions.

Research Question # 1

After four years, what is the relationship between MAP scores and performance on widely used assessments commonly considered to be strong correlates or direct measures of academic success?

The standardized criterion measures used as dependent measures in the correlational analyses included the following (abbreviations as used in the Tables are noted in parentheses after each test name):

The Weschler Intelligence Scale-Revised Full Scale Intelligence Quotient (WISC-R Full Scale)

The Weschler Intelligence Scale-Revised Performance Intelligence Quotient (WISC-R Perform)

The Weschler Intelligence Scale-Revised Verbal Intelligence Quotient (WISC-R Verbal)

The Berry Developmental Test of Visual Motor Integration (Beery Visual Motor, VMI,or Beery)

The Bruininks - Oseretsky Test of Motor Proficiency (Bruininks or Bruin) This scale is further subdivided into:
  - Bruininks Total Score (Bruin Total)
  - Bruininks Gross Motor Score (Bruin Gross)
  - Bruininks Fine Motor Score (Bruin Fine)

Table 1
Distribution of Total Scores On MAP In Sample Studied

Total Map Score (Reported in percentiles)	# of Children Receiving Score	Percentage Of Sample Receiving Score	Cumulative Percent
0 - 5%	15	4.4%	4.4%
6 - 10%	26	7.7%	12.1%
11 - 15%	14	4.1%	16.2%
16 - 20%	26	7.7%	23.9%
21 - 25%	0	0	23.9%
26 - 30%	29	8.6%	32.5%
31 - 35%	24	7.1%	39.6%
36 - 40%	21	6.2%	45.8%
41 - 45%	0	0	45.8%
46 - 50%	22	6.5%	52.3%
51 - 55%	36	10.7%	63.0%
56 - 60%	0	0	63.0%
61 - 65%	42	12.4%	75.4%
66 - 70%	0	0	75.4%
71 - 75%	44	13.0%	88.4%
76 - 80%	0	0	88.4%
81 - 85%	18	5.3%	93.7%
86 - 90%	0	0	93.7%
91 - 95%	15	4.4%	98.1%
96 - 100%	6	1.8%	99.9%
Column Totals	338	99.9%	99.9%

Table 2
Description of Performance of Sample On Standardized Criterion Measures

	Mean Score of Sample	Standard Deviation of Sample	Range of Sample	Mean Derived Score of Norming Population	Standard Deviation of Norming Population
Beery Visual Motor	9.37	3.08	1 - 19	10	3
WISC-R Verbal	114.09	16.60	47 - 153	100	15
WISC-R Performance	113.05	14.92	45 - 147	100	15
WISC-R Full Scale	115.11	16.25	41 - 152	100	15
Bruininks Motor	59.14	11.17	24 - 75	50	10
Goodenough Drawing	98.60	16.52	58 - 151	100	15
Woodcock Reading	105.27	13.52	65 - 135	100	15
Woodcock Math	105.46	14.46	65 - 135	100	15
Woodcock Language	107.06	14.71	65 - 135	100	15
Walker Parent Total	53.92	10.80	39 - 99	*	*
Walker Teacher Total	50.85	10.35	41 - 99	*	*

* The Walker has point scores which are translated into T-score distributions. The Walker Manual notes that T-scores between 40 - 60 are normal, and a score above 60 has been selected as a cutoff point suggesting referral for further evaluation and/or testing. Means and standard deviations for norming population are not reported in manual.

The Goodenough - Harris Drawing Test (Goodenough Drawing or Goodenough)

The Woodcock-Johnson Psychoeducational Battery -Reading (Woodcock Reading)

The Woodcock-Johnson Psychoeducational Battery -Mathematics (Woodcock Math)

The Woodcock-Johnson Psychoeducational Battery - Language (Woodcock Language)

The Walker Problem Behavior Identification Checklist: Parent form (Walker Parent or Walker P)

The Walker Problem Behavior Identification Checklist : Teacher form (Walker Teacher or Walker T)

Teachers' Rating of child's overall achievement (Teacher Overall Rating)

Pearson Product Moment correlations were calculated for each of the twelve criterion measures with the MAP Total Score, each MAP subtest (called Index), and each item. The bivariate correlations for the MAP total score and MAP subtests, and related p values are reported in Table 3. The results of the correlations for each MAP test item and each of the major twelve criterion measures are reported in Tables 4 - 8. Tables 4 - 8 are noted in the order in which the subtests appear in the MAP: Table 4 - Foundations Index; Table 5 - Coordination Index; Table 6 - Verbal Index; Table 7 - Non-verbal Index; and Table 8 - Complex Tasks Index. The items are specified in the subtest in which they appear, in the order of administration.

The following text first discusses the interpretation of predictive correlation coefficients, then summarizes findings and discussion for the MAP Total Score, and then explores data and discussion for each MAP subtest (Index).

Interpretation of Predictive Correlation Coefficients

The presentation of the correlative results of a predictive validity study should be a straightforward task, involving the reporting of varying amounts of data, generated through well-defined statistical means. However, the interpretation of the results of predictive validity research is exceedingly treacherous, the entire area being littered with sources of invalidity waiting to trip up the unwary author. Nowhere are these difficulties more in evidence, than in the interpretation of correlation coefficients in predictive validity studies.

Of particular note is the subtle, almost subliminal, interpretation embedded in the use of descriptive modifiers for numerical correlation coefficients. Adjectives such as "moderately high", "fairly high", "useful", "acceptable", "firm", "above average", "good", "respectable", "efficient", etc., are not only entirely subjective in nature (and, therefore, evocative of different value judgements in different readers) but also have been used in an almost completely haphazard way in published predictive validity research. A firm effort has therefore been made in this report to eliminate, as much as possible, arbitrary interpretation of results. Results are discussed in terms of significance (i.e. statistical p values described by significant, highly significant, etc.) instead of descriptive modifiers.

The sine que non of predictive validity is a time lapse between the determination of the predictor variable and the observed performance on a criterion measure. The length of this time lapse is a unique and inseparable component of the correlation coefficient describing the relationship between the predictor variable and the criterion measure. In other words, one cannot rationally evaluate the practical value of a correlation coefficient in predictive validity research without taking into account the time lapse involved (Sax, 1968). Simplistic guides, such as those presented by Darlington, (1975) Borg and Gall, (1983) and Fox, (1969) which assign terms such as "high" or "moderate" to absolute numerical correlation coefficients are, therefore, meaningless in predictive validity research. These authors generally site specific correlational values and assign a descriptive modifier to each value. For example Fox (1969) states the following categorizations of absolute values of correlations: "low, .50; moderate, .50 to .70; high, .70 to .86; very high, above .86. Many other authors also offer guidelines such as these. It is essential to note that these guidelines refer to correlation coefficients for concurrent validity studies only. Since the time lapse which is inherently a part of all predictive research is an inseparable component of the correlation coefficient (over time the coefficient will go to 0), these guidelines for concurrent validity coefficients are not applicable to predictive validity correlations coefficients.

Table 3
Correlations Between All MAP Indices, MAP Total Score, MAP Subtest Scores And Twelve Criterion Measures

CRITERION MEASURES	MAP SUBTESTS AND MAP TOTAL SCORE					
	FOUNDATIONS INDEX	COORDINATION INDEX	VERBAL INDEX	NON-VERBAL INDEX	COMPLEX TASKS INDEX	MAP TOTAL SCORE
BEERY VISUAL MOTOR	.10 p=.03	.11 p=.03	.07 p=.09	.21 p=.001	.19 p=.001	.21 p=.001
WISC-R VERBAL	.28 p=.001	.27 p=.001	.32 p=.001	.25 p=.001	.32 p=.001	.45 p=.001
WISC-R PERFORMANCE	.33 p=.001	.33 p=.001	.19 p=.001	.33 p=.001	.39 p=.001	.47 p=.001
WISC-R FULL SCALE	.33 p=.001	.32 p=.001	.29 p=.001	.31 p=.001	.38 p=.001	.50 p=.001
BRUININKS MOTOR	.28 p=.001	.27 p=.001	.23 p=.001	.20 p=.001	.28 p=.001	.39 p=.001
GOODENOUGH DRAWING	.08 p=.07	.13 p=.007	.06 p=.135	.14 p=.005	.20 p=.001	.19 p=.001
WOODCOCK READING	.18 p=.001	.26 p=.001	.30 p=.001	.22 p=.001	.25 p=.001	.36 p=.001
WOODCOCK MATH	.23 p=.001	.23 p=.001	.28 p=..001	.19 p=.001	.30 p=.001	.38 p=.001
WOODCOCK LANGUAGE	.21 p=.001	.25 p=.001	.25 p=.001	.22 p=.001	.21 p=.001	.35 p=.001
WALKER TOTAL PARENT	.08 p=.07	.08 p=.07	.15 p=.002	.04 p=.237	.13 p=.008	.17 p=.001
WALKER TOTAL TEACHER	.08 p=.08	.09 p=.06	.15 p=.005	.04 p=.259	.09 p=.05	.13 p=.008
TEACHER OVERALL RATING	.12 p=.079	.17 p=.002	.18 p=.001	.13 p=.01	.18 p=.001	.23 p=.001

Table 4
Correlations Between MAP Foundations Index Items And Twelve Criterion Measures

CRITERION MEASURES	VERTICAL WRITING	ROMBERG	STAMP	HAND-NOSE	STEREOGNOSIS	FINGER LOCALIZATION	WALK LINE	STEPPING	SUPINE FLEXION	KNEEL-STAND
	MAP ITEMS IN FOUNDATIONS INDEX									
BEERY VISUAL MOTOR	.01 p=.435	.05 p=.204	.04 p=.212	.09 p=.053	.05 p=.159	.07 p=.106	.09 p=.050	.07 p=.112	.13 p=.01	.05 p=.182
WISC-R VERBAL	.02 p=.352	.19 p=.001	.24 p=.001	.21 p=.001	.23 p=.001	.09 p=.054	-.07 p=.086	.17 p=.001	.10 p=.032	.12 p=.011
WISC-R PERFORMANCE	.11 p=.022	.18 p=.001	.19 p=.001	.16 p=.002	.24 p=.001	.15 p=.004	-.01 p=.444	.16 p=.001	.11 p=.020	.15 p=.003
WISC-R FULL SCALE	.07 p=.110	.19 p=.001	.24 p=.001	.20 p=.001	.26 p=.001	.12 p=.012	-.05 p=.191	.18 p=.001	.11 p=.020	.15 p=.003
BRUININKS MOTOR	.02 p=.380	.22 p=.001	.20 p=.001	.16 p=.002	.17 p=.001	.01 p=.441	.01 p=.427	.14 p=.005	.19 p=.001	.11 p=.024
GOODENOUGH DRAWING	.02 p=.364	.00 p=.489	-.03 p=.271	.07 p=.102	-.01 p=.454	.04 p=.222	.00 p=.493	.08 p=.063	.09 p=.04	.08 p=.070
WOODCOCK READING	.03 p=.284	.09 p=.056	.16 p=.002	.14 p=.006	.10 p=.038	.06 p=.147	-.00 p=.467	.13 p=.008	.11 p=.021	.02 p=.381
WOODCOCK MATH	.00 p=.494	.15 p=.003	.16 p=.001	.13 p=.007	.22 p=.001	.07 p=.100	-.03 p=.306	.17 p=.001	.11 p=.022	.05 p=.161
WOODCOCK LANGUAGE	.03 p=.273	.14 p=.005	.15 p=.002	.12 p=.012	.17 p=.001	.08 p=.083	-.01 p=.453	.11 p=.022	.12 p=.011	.12 p=.011
WALKER TOTAL PARENT	.01 p=.431	.10 p=.032	.00 p=.496	.09 p=.054	.15 p=.003	.06 p=.141	.13 p=.009	.12 p=.368	.03 p=.263	.08 p=.073
WALKER TOTAL TEACHER	.03 p=.285	.15 p=.003	.11 p=.025	.01 p=.406	.08 p=.067	.00 p=.471	.05 p=.181	.04 p=.264	.02 p=.349	.00 p=.479
TEACHER OVERALL RATING	-.04 p=.227	.16 p=.002	.06 p=.132	.09 p=.059	.08 p=.083	.09 p=.059	.05 p=.185	.07 p=.116	.07 p=.118	-.05 p=.189

Table 5
Correlations Between MAP Coordination Index Items And Twelve Criterion Measures

CRITERION **MEASURES**	MAP ITEMS IN COORDINATION INDEX						
	TOWER	CAGE	STAMP	TONGUE MOVEMENT	WALK LINE	ARTICU-LATION	VERTICAL WRITING
BEERY VISUAL MOTOR	.03 p=.319	.13 p=.008	.04 p=.212	.04 p=.242	-.09 p=.050	.20 p=.001	.01 p=.435
WISC-R VERBAL	.16 p=.002	.14 p=.005	.24 p=.001	.05 p=.164	-.07 p=.086	.34 p=.001	.02 p=.352
WISC-R PERFORMANCE	.15 p=.003	.18 p=.001	.19 p=.001	.12 p=.013	-.01 p=.444	.29 p=.001	.11 p=.022
WISC-R FULL SCALE	.17 p=.001	.17 p=.001	.24 p=.001	.12 p=.012	-.05 p=.191	.34 p=.001	.07 p=.110
BRUININKS MOTOR	.20 p=.001	.21 p=.001	.20 p=.001	.01 p=.441	.01 p=.427	.21 p=.001	.02 p=.380
GOODENOUGH DRAWING	.02 p=.352	.13 p=.007	-.03 p=.271	.04 p=.222	.00 p=.493	.18 p=.001	02 p=.364
WOODCOCK READING	.15 p=.003	.19 p=.001	.16 p=.002	.07 p=.110	-.00 p=.467	.24 p=.001	.03 p=.284
WOODCOCK MATH	.11 p=.019	.19 p=.001	.16 p=.001	.01 p=.439	-.03 p=.306	.29 p=.001	.00 p=.494
WOODCOCK LANGUAGE	.12 p=.014	.15 p=.003	.15 p=.002	.07 p=.093	-.01 p=.453	.26 p=.001	.03 p=.273
WALKER TOTAL PARENT	.14 p=.006	.18 p=.001	.00 p=.496	.13 p=.343	.13 p=.009	.13 p=.009	.01 p=.431
WALKER TOTAL TEACHER	.02 p=.393	.07 p=.115	.11 p=.025	.09 p=.042	.05 p=.181	.05 p=.202	.03 p=.285
TEACHER OVERALL RATING	.05 p=.182	.21 p=.001	.06 p=.132	.08 p=.081	.05 p=.185	.14 p=.005	- .04 p=.227

Table 6
Correlations Between MAP Verbal Index Items And Twelve Criterion Measures

CRITERION MEASURES	MAP ITEMS IN VERBAL INDEX			
	DIGIT REPETITION	GENERAL INFORMATION	FOLLOW DIRECTIONS	SENTENCE REPETITION
BEERY VISUAL MOTOR	.13 p=.010	.05 p=.193	-.02 p=.348	.04 p=.237
WISC-R VERBAL	.31 p=.001	.20 p=.001	.11 p=.021	.22 p=.001
WISC-R PERFORMANCE	.20 p=.001	.14 p=.004	.01 p=.448	.14 p=.004
WISC-R FULL SCALE	.29 p=.001	.19 p=.001	.07 p=.101	.20 p=.001
BRUININKS MOTOR	.21 p=.001	.17 p=..001	.02 p=.373	.20 p=.001
GOODENOUGH DRAWING	.13 p=.011	-.00 p=.490	.01 p=.431	.03 p=.297
WOODCOCK READING	.30 p=.001	.20 p=.001	.08 p=.063	.22 p=.001
WOODCOCK MATH	.32 p=.001	.15 p=.002	.06 p=.125	.20 p=.001
WOODCOCK LANGUAGE	.29 p=.001	.15 p=.002	.06 p=.120	.16 p=.002
WALKER TOTAL PARENT	.15 p=.004	.12 p=.013	.06 p=.118	.08 p=.079
WALKER TOTAL TEACHER	.10 p=.041	.13 p=.008	.06 p=.131	.09 p=.056
TEACHER OVERALL RATING	.18 p=.001	.07 p=.114	.11 p=.020	.11 p=.027

Table 7
Correlations Between MAP Non Verbal Index Items And Twelve Criterion Measures

CRITERION MEASURES	MAP ITEMS IN NONVERBAL INDEX				
	SEQUENCING	BLOCK TAPPING	OBJECT MEMORY	FIGURE GROUND	PUZZLES
BEERY VISUAL MOTOR	.05 p=.189	.12 p=.012	.16 p=.126	.13 p=.008	.19 p=.001
WISC-R VERBAL	.05 p=.179	.17 p=.001	-.04 p=.251	.23 p=.001	.29 p=.001
WISC-R PERFORMANCE	.08 p=.090	.23 p=.001	.05 p=.195	.25 p=.001	.31 p=.001
WISC-R FULL SCALE	.07 p=.090	.21 p=.001	-.01 p=.490	.26 p=.001	.33 p=.001
BRUININKS MOTOR	.06 p=.121	.08 p=.066	.06 p=.047	.16 p=.002	.18 p=.001
GOODENOUGH DRAWING	.06 p=.131	.11 p=.022	-.01 p=.397	-.05 p=.186	.19 p=.001
WOODCOCK READING	.01 p=.403	.17 p=.001	.04 p=.210	.10 p=.027	.28 p=.001
WOODCOCK MATH	-.02 p=.326	.16 p=.002	.01 p=.424	.15 p=.003	.23 p=.001
WOODCOCK LANGUAGE	.05 p=.184	.19 p=.001	.08 p=.066	.13 p=.008	.24 p=.001
WALKER TOTAL PARENT	.06 p=.153	.04 p=.240	.07 p=.086	.09 p=.046	.12 p=.015
WALKER TOTAL TEACHER	.01 p=.454	.01 p=.442	.04 p=.256	.05 p=.344	.06 p=.158
TEACHER OVERALL RATING	.01 p=.411	.15 p=.003	.01 p=.395	.08 p=.080	.13 p=.010

Table 8
Correlations Between MAP Complex Tasks Index Items And Twelve Criterion Measures

CRITERION MEASURES	MAP ITEMS IN COMPLEX TASKS INDEX			
	DRAW A PERSON	MAZE	BLOCK DESIGNS	IMITATION POSTURES
BEERY VISUAL MOTOR	.10 p=.03	.17 p=.001	.14 p=.005	.01 p=.423
WISC-R VERBAL	.21 p=.001	.19 p=.001	.22 p=.001	.12 p=.017
WISC-R PERFORMANCE	.22 p=.001	.28 p=.001	.28 p=.001	.09 p=.043
WISC-R FULL SCALE	.23 p=.001	.25 p=.001	.27 p=.001	.27 p=.001
BRUININKS MOTOR	.16 p=.002	.14 p=.006	.24 p=.001	.24 p=.001
GOODENOUGH DRAWING	.12 p=.02	.11 p=.026	.20 p=.001	.20 p=.001
WOODCOCK READING	.13 p=.008	.19 p=.001	.18 p=.001	.18 p=.001
WOODCOCK MATH	.19 p=.001	.18 p=.001	.25 p=.001	.25 p=.001
WOODCOCK LANGUAGE	.11 p=.019	.17 p=.001	.16 p=.002	.16 p=.002
WALKER TOTAL PARENT	.03 p=.295	.08 p=.063	.13 p=.010	.13 p=.010
WALKER TOTAL TEACHER	.01 p=.426	.08 p=.080	.08 p=.072	.08 p=.072
TEACHER OVERALL RATING	.04 p=.246	.17 p=.001	.14 p=.006	.06 p=.147

Although the integral nature of the time lapse in predictive validity research seems obvious, most authors do not discuss it as a separate issue. For example, Borg and Gall (1983) fail to even mention it in their section "Interpreting Magnitude of Correlation Coefficients". In contrast, Sax (1968) states, "to evaluate predictive validity coefficients, one would need to know the amount of time elapsing between the administration of the predictor and the criterion. In this sense, _no test has just one validity coefficient_" (italics in original).

Given the necessity to consider the time lapse as an integral factor in the correlation coefficient, how does one proceed in a reasonable way to assign interpretations of practical value to this type of predictive validity data? Darlington (1975) suggests, "A somewhat more accurate way to interpret the size of a particular correlation coefficient is to gain some familiarity with the sizes of correlations typically observed in the same area of research. A correlation may then be described by such phrases as 'one of the highest correlations ever observed in this type of research' or 'typical of other correlations in this area'".

This approach is of particular importance given the fact that, theoretically, the correlation will approach zero as the time lapse increases (Bowerman and O'Connell, 1979; Box and Jenkins, 1976). The practical value of a predictive validity correlation of .50 is completely dependent on whether the time lapse is one month, one year, or four years.

There are several other factors which tend to increase the practical value of the correlation coefficients reported in this study. In general these include the absence of sources of external invalidity. In addition, as described by Borg and Gall (1983) favorable selection ratio enhances the practical value of correlation coefficients, as does a low tendency for "natural selection" to accurately predict the criterion measure. However, it is still felt that Darlington's suggestion for interpretation is most appropriate, that is, to compare present results to results documented in research which is similar in length of time and variables assessed.

Table 9 summarizes previous research findings for studies similar in length and content. It can be seen from this table that the correlation coefficients of the present study compare favorably to those of predictive studies having a similar design and time lapse.

1. Analysis of Correlational Data for the MAP Total Score

As indicated in Table 3, the correlation between the MAP and 11 of the 12 criterion measures was significant at the .001 level,and the correlation with the remaining criterion measure was significant at the .01 level. Of particular note is the highly significant correlation between the MAP Total score and the WISC-R measures (r = .50 to .45) and between the MAP Total Score and the Woodcock Math, Reading, and Language measures (r = .38 to .35). The correlation of the MAP and the Bruininks was also highly significant (r = .39) In effect, these correlations indicate that even though the MAP was administered four years previous to the administration of the criterion measures, the correlations were significantly higher than would be expected simply on the basis of chance.

In order to get an idea of how well the MAP predicted inter-correlations between the criterion measures were computed and compared to the correlations of the MAP Total Score with the criterion measures. These data are summarized in Table 10. As can be seen in from this table, the intercorrelations between these standardized measures is not as high as might be hypothesized would result from these measures being administered at the same time. It is interesting that the WISC-R, and Woodcock measures (which are felt by many researchers to be highly related) correlate at the .61 to .68 level. When viewed in this light, the correlation of the MAP to the WISC-R of .50 seems high, considering it was administered four years previously.

Another interesting note from this table is that the Beery and the Goodenough correlate most highly with each other (r = .45), but do not correlate well with the other measures. The MAP scores tended not to correlate as highly with the Beery and Goodenough as they did with the WISC-R and Woodcock measures. It seems likely from these data the the MAP, WISC-R, and Woodcock covary, more than do the MAP and the Beery or Goodenough. Since the focus of this study was the MAP, no attempt will be made to fully discuss the interrelationships of the criterion measures; however for purposes of valuating the correlations between the MAP and

Table 9

Summary of Correlational Studies by Descriptive Modifiers

Author/Date	Interval Studied	Correlation Coefficient	Descriptive Modifier
Nichta et al (1982)	3.5 years	.71	Useful
Dziuban & Mealor (1982)	4 years	.54	Moderate But Firm
Lindeman et al (1984)	3 years	.67	Moderate
Colligan (1979)	2 years	.53	Considerable Potential
Klein (1977)	1 - 2 yrs.	.57 - .70	Efficient
Keough & Smith (1967)	7 years	.30 - .41	Lend substantial support for use
Obrzut et al (1981)	1 year	.30 - .41	Useful
Davies (1980)	1 1/2 years	.60 - .80	Substantial Relationship
Flook et al (1977)	2 months	.67 - .80	Very High
Telegdy (1975)	1 year	.58 - .73	Powerful Predictors
Brown (1976)	6 months	.33 - .63	Moderately High
Wallbrown et al (1975)	1 year	.55 - .70	Statistically Significant
Densan et al (1974)	6 months	.26 - .66	Valid Predictor
Flynn & Flynn (1978)	2 years	.31 - .38	Significant Relationship

Table 10

Intercorrelations of MAP Total Score and the Standardized Criterion Measures

	Beery	IQ	Bruininks	Goodenough	Reading	Math	Language	Walker P	Walker T	Map Total Test
Beery Visual Motor	1.00	.48	.35	.45	.40	.38	.35	.18	.17	.21
WISC-R Full Scale		1.00	.46	.36	.64	.68	.61	.30	.17	.50
Bruininks Motor			1.00	.22	.39	.44	.38	.23	.13	.39
Goodenough Drawing				1.00	.33	.24	.33	.01	.04	.19
Woodcock Reading					1.00	.65	.81	.27	.25	.36
Woodcock Math						1.00	.68	.30	.15	.38
Woodcock Language							1.00	.29	.21	.35
Walker Total Parent								1.00	.36	.18
Walker Total Teacher									1.00	.13
MAP Total Score										1.00

the criterion measures, the intercorrelations provide an interesting comparison.

In terms of predictive validity, it can be concluded that the administration of the MAP at a preschool level can predict intelligence and achievement four years later at a level typically observed in similar research.

All MAP indices except for the Foundations Index correlated at a highly significant level with the criterion measures and even the Foundations Index had a significant correlation with seven of the 12 criterion measures. However, no MAP index had a better set of correlations than the MAP Total score. This indicates that the MAP Total Score is a better predictor of future performance than any of the specific subtests of the MAP.

2. Analysis of Correlational Data for each of the Five Subtests of the MAP

Analysis of each of five subtests, or indices of the MAP occurs in the order of administration as follows: a) Foundations Index; b) Coordination Index; c) Verbal Index; d) Non-Verbal Index; e) Complex Tasks Index.

a) Foundations Index

The FI correlates most highly with the IQ measures (with correlations ranging from .28 to .33), and with the Bruininks Motor Scale (r = .28). In addition, the Woodcock Reading, Math and Language scores were significant at the .001 level (r = .18 to .23). None of the other criterion measures correlate with the FI at a significant level.

It is not surprising that the Bruininks and MAP FI correlate, since the FI includes the neurological items, many of which have a motor base. However, the highest correlations (the FI with IQ and achievement) are surprising. It is possible that the early correlates of what is measured by the WISC-R and Woodcock have a stronger neurological base than previously demonstrated by other research.

b) Coordination Index

The CI correlated most highly with Performance IQ, and Full Scale IQ (r = .33 and .32 respectively). In addition, the Bruininks, Verbal IQ, Woodcock Reading, Math and Language tests also correlate at a highly significant level (r= .23 to .27).

It is logical for the CI and the Bruininks to correlate highly since they are both measures of motor ability. As with the FI a large correlation between CI and IQ was unexpected and may speak to early antecedents of what is measured by the WISC-R. The relatively low correlations with the Beery and Goodenough were surprising since the CI measures in part, fine motor performance.

c) Verbal Index

The highest correlations with the VI were the Verbal IQ and the Woodcock Reading scores (r = .32 and .30 respectively). In addition significant correlations were found between the CI and Full Scale IQ, Woodcock Math and Language,and the Bruininks (r = .25 to .29),and the Performance IQ,Teacher Overall Rating, and Walker Total parent and Walker Total teacher rating at a slightly lower level, still highly significant (r = .15 to .19).

It is interesting to note that the Verbal Index correlated at a highly significant level with the Verbal IQ measure, and with reading. This indicates that there is a strong linear relationship between the two measures. The Verbal Index most accurately predicts Verbal IQ, when compared to the other criterions measured.

d) Non verbal Index

The highest correlations occurred between the NVI and the Performance and Full Scale IQ measures (r = .33 and .31). In addition, highly significant correlations were noted between the NVI and the Verbal IQ, Woodcock Reading, Math, and Language, the Beery, the Bruininks, and the Goodenough (r = .14 to .25).

It is interesting to note that the NVI and CTI (discussed below) are the only subtests with highly significant correlations to the Beery and Goodenough. These two so called "visual-motor" measures are frequently included in the assessment of children for dysfunction. They appear to correlate more highly with measures of nonverbal cognitive function, than with fine motor skills.

e) Complex Tasks Index

The correlations between the CTI and the Performance and Full Scale IQ measures were highly significant ([illegible] = .39 and .38). Other highly significant correlations with the CTI included: Verbal IQ, Woodcock Math, Bruininks, Woodcock Reading, Woodcock Language, Goodenough, Beery, and Teacher Overall Rating (r = .18 to .32).

The items which appear in the CTI were put into that particular subtest because they measure behaviors which represent "combined" abilities, such as visual and motor, or cognitive and movement. Considering this, it is interesting that Performance IQ correlates more highly than Verbal IQ, and that the correlation between the Bruininks, a motor scale, is higher than that of the Woodcock Reading and Language measures, which are non-motor in nature. This may suggest that the items measured in this index are more predictive of motor and performance later, than they are of cognitive skills.

3. Analysis of Correlational Data for each of the Twenty Seven Items of the MAP

Tables 4 - 8 note the correlation of each of the specific MAP items with the same twelve criterion measures described above in Table 3. However, for ease of interpretation, tables 11 - 15 were constructed which rank order the criterions to which each item correlated at a .01 level or better. There is a separate Table for each of the five subtests of the MAP denoting each item from the test. In addition to the twelve criterion measures described previously, other measures for which abbreviations are noted in Figure 1, were ranked in Tables 11 - 15.

Table 11

Correlations and p Values for Items
in the MAP Foundations Index

Correlations Significant at .01 or better

STEREOGNOSIS			FINGER LOCALIZATION			VERTICAL WRITING	HAND-NOSE			ROMBERG		
FULL IQ	.26	p=.001	DAP #1	.24	p=.001	SEE	FULL IQ	.20	p=.001	H PHYS T	.28	p=.001
PERFORM IQ	.24	p=.001	PERFORM IQ	.15	p=.004	COORDINATION	VERBAL IQ	.21	p=.001	BRUIN TOTAL	.22	p=.001
VERBAL IQ	.23	p=.001	H SOC C	.14	p=.004	INDEX	BRUIN GROSS	.20	p=.001	DAP #12	.21	p=.002
WJ MATH	.22	p=.001	FULL IQ	.12	p=.012		BRUIN TOTAL	.18	p=.001	HMAT ATTACH	.20	p=.013
WJ LANG	.18	p=.001					WJ MATH	.17	p=.001	FULL IQ	.19	p=.001
W DISTRAC P	.18	p=.001					PERFORM IQ	.16	p=.002	BRUIN GROSS	.19	p=.001
BRUIN TOTAL	.17	p=.001					WJ LANG	.16	p=.001	VERBAL IQ	.18	p=.001
DAP #9	.17	p=.009					T MEMORY	.15	p=.004	DAP #9	.18	p=.008
H PEER T	.16	p=.004					T BAL/COOR	.15	p=.004	T LANGUAGE	.18	p=.001
H PEER C	.15	p=.003					WJ READ	.14	p=.006	PERFORM IQ	.17	p=.001
W TOTAL P	.15	p=.003					H PEER T	.14	p=.011	W PEER T	.16	p=.002
H COG T	.13	p=.009					T PHYS ACT	.14	p=.005	T MATH	.16	p=.002
BRUIN GROSS	.12	p=.016					T DEXTER	.13	p=.010	T OVERALL	.16	p=.002
T VERBAL	.12	p=.015					W ACT OUT P	.12	p=.013	WJ MATH	.15	p=.003
										H COG T	.15	p=.003
										W TOT T	.15	p=.003
										BRUIN FINE	.14	p=.006
										WJ LANGUAGE	.14	p=.004
										T FOL DIR	.13	p=.011
										H SOC T	.12	p=.018
										T READING	.12	p=.013

# WISC SUBTESTS Significant at .01 level	STEREOGNOSIS	FINGER LOCALIZATION	VERTICAL WRITING	HAND-NOSE	ROMBERG
	10	3	1	7	7

Table 11
Page 2

	STEPPING			WALKS LINE	SUPINE FLEXION			KNEEL-STAND			STAMP
	H PHYS C	.20	p=.001	SEE COORDINATION INDEX	H CONDUCT C	.22	p=.001	BRUIN GROSS	.19	p=.001	SEE COORDINATION INDEX
	FULL IQ	.18	p=.001		BRUIN TOTAL	.19	p=.001	PERFORM IQ	.15	p=.003	
	VERBAL IQ	.17	p=.001		BRUIN GROSS	.16	p=.001	FULL IQ	.15	p=.003	
	WJ MATH	.17	p=.001		BEERY	.15	p=.004	H SOC C	.14	p=.006	
	PERFORM IQ	.16	p=.001		GOODENOUGH	.15	p=.003	BRUIN TOTAL	.13	p=.008	
	DAP 011	.16	p=.012		H PHYS C	.15	p=.004	T DEXTER	.13	p=.01	
	H COG T	.16	p=.003		H PHYS T	.15	p=.005	VERBAL IQ	.12	p=.011	
	H APPEAR C	.16	p=.012		BRUIN FINE	.14	p=.001				
	H COG C	.15	p=.003		WJ LANGUAGE	.14	p=.004				
	H SOC C	.15	p=.003		T BAL COOR	.13	p=.008				
	BRUIN TOTAL	.14	p=.005		T PHYS ACT	.13	p=.010				
	WJ READ	.13	p=.008		T DEXTER	.13	p=.011				
	H PHYS T	.13	p=.017		WJ READING	.12	p=.012				
					WJ MATH	.12	p=.014				
# WISC SUBTESTS Significant at .01 level	5				3			6			10

Table 12

Correlations and Values For Items in the MAP Coordination Index

Correlations Significant at .01 or better

TOWER

T LANGUAGE	.21	p=.001
BRUIN TOTAL	.20	p=.001
W PHYS T	.20	p=.001
T BAL/COORD	.19	p=.001
FULL IQ	.17	p=.001
VERBAL IQ	.16	p=.002
PERFORM IQ	.15	p=.003
WJ READ	.15	p=.003
T WRIT EXP	.15	p=.004
T MEMORY	.13	p=.008
WJ LANG	.12	p=.014
T PHYS ACT	.12	p=.013
WJ MATH	.11	p=.019

CAGE

T MATH	.26	p=.001
T ATTN SPAN	.26	p=.001
W PEER P	.24	p=.001
BRUIN TOTAL	.23	p=.001
DAP 812	.22	p=.001
WJ READ	.21	p=.001
WJ MATH	.21	p=.001
T READING	.21	p=.001
T OVERALL	.21	p=.001
BRUIN FINE	.20	p=.001
GOODENOUGH	.20	p=.001
T BAL/COORD	.19	p=.001
PERFORM IQ	.18	p=.001
W TOTAL P	.18	p=.001
T HAND WRTG	.18	p=.001
T WRT EXP	.18	p=.001
FULL IQ	.17	p=.001
BRUIN GROSS	.17	p=.001
WJ LANG	.16	p=.001
T LANGUAGE	.16	p=.002
T FOL DIR	.16	p=.003
W IMMATURE P	.15	p=.002
T PROB SOLV	.15	p=.004
T DEXTER	.15	p=.003
VERBAL IQ	.14	p=.005
W PEER T	.14	p=.011
T VOCAB	.14	p=.006
BEERY	.13	p=.008
W DISTRACT P	.13	p=.011
T MOTIV	.13	p=.008
T MEMORY	.12	p=.017

VERTICAL WRITING

Bruin Fine	.12	p=.012

WALKS LINE

BEERY	.13	p=.008
W ACT OUT P	.13	p=.009
W TOTAL P	.13	p=.009
BRUIN GROSS	.12	p=.014
WJ MATH	.12	p=.017

WISC SUBTESTS
Significant at
.01 level

6+ ALL OTHERS <.05

Table 12
Page 2

STAMP

VERBAL IQ	.24	p=.001
FULL IQ	.24	p=.001
BRUIN TOTAL	.21	p=.001
BRUIN GROSS	.20	p=.001
PERFORM IQ	.19	p=.001
WJ MATH	.18	p=.001
WJ READ	.17	p=.001
WJ LANG	.17	p=.001
BRUIN FINE	.13	p=.01
W WITHDRAW T	.12	p=.017
W IMMATURE T	.12	p=.014

TONGUE MOVEMENT

DAP #12	.19	p=.005
DAP #5	.17	p=.009
PERFORM IQ	.12	p=.013
GOODENOUGH	.12	p=.013
T HAND WTG	.12	p=.018

ARTICULATION

VERBAL IQ	.34	p=.001
FULL IQ	.34	p=.001
PERFORM IQ	.29	p=.001
WJ MATH	.29	p=.001
WJ LANG	.26	p=.001
WJ READ	.24	p=.001
BRUIN TOTAL	.21	p=.001
T VERBAL	.21	p=.001
BEERY	.20	p=.001
BRUIN GROSS	.20	p=.001
H PHYS T	.19	p=.001
GOODENOUGH	.18	p=.001
T LANGUAGE	.17	p=.001
T HAND WTG	.16	p=.003
T WTN EXP	.16	p=.002
H SOC T	.15	p=.004
T OVERALL	.14	p=.005
H COG T	.13	p=.008
H PHYS C	.13	p=.011
H SOC C	.13	p=.008
W WITHDRAW T	.13	p=.009
W TOTAL P	.13	p=.009
T MATH	.13	p=.009
T VOCAB	.13	p=.011
W ACT OUT P	.12	p=.013
W DISTRACT P	.12	p=.013

	STAMP	TONGUE MOVEMENT	ARTICULATION
# WISC SUBTESTS Significant at .01 level	10	3	10

Table 13

Correlations and p Values for Items In the MAP Verbal Index

	GENERAL INFORMATION			FOLLOW DIRECTIONS	SENTENCE REPETITION			DIGIT REPETITION		
	VERBAL IQ	.20	p=.001	0	WJ READ	.23	p=.001	WJ MATH	.32	p=.001
	WJ READ	.20	p=.001		VERBAL IQ	.22	p=.001	VERBAL IQ	.31	p=.001
	FULL IQ	.19	p=.001		WJ MATH	.21	p=.001	WJ READ	.30	p=.001
	BRUIN TOTAL	.17	p=.001		FULL IQ	.20	p=.001	WJ LANG	.30	p=.001
	WJ LANG	.17	p=.001		BRUIN TOTAL	.20	p=.001	TOTAL IQ	.29	p=.001
	W WITHDRAW P	.17	p=.001		BRUIN GROSS	.17	p=.001	BRUIN TOTAL	.21	p=.001
	WJ MATH	.15	p=.002		WJ LANG	.17	p=.001	T LANGUAGE	.21	p=.001
	PERFORM IQ	.14	p=.004		DAP 012	.16	p=.013	PERFORM IQ	.20	p=.001
	W ACT OUT T	.14	p=.007		PERFORM IQ	.14	p=.004	H COG T	.20	p=.001
	BRUIN GROSS	.13	p=.008		T READING	.14	p=.007	T READ	.20	p=.001
	H COG T	.13	p=.010		T LANGUAGE	.13	p=.010	T WRIT EXP	.20	p=.001
	H SOC T	.13	p=.012		BRUIN FINE	.12	p=.013	T MATH	.18	p=.001
	W TOT T	.13	p=.008					T OVERALL	.18	p=.001
	W IMMAT T	.12	p=.013					T VERBAL	.17	p=.001
	W TOT P	.12	p=.013					T MEM	.17	p=.001
								W ACT OUT P	.15	p=.002
								W TOT P	.15	p=.004
								T MOTIV	.15	p=.004
								BRUIN FINE	.14	p=.006
								BEERY	.13	p=.010
								GOODENOUGH	.13	p=.007
								BRUIN GROSS	.12	p=.016
								H COG C	.12	p=.011
								W ACT OUT T	.12	p=.018
								W PEER P	.12	p=.013
								T VOCAB	.12	p=.018
								T PROB SOLV	.12	p=.013
								T ATTN SPAN	.12	p=.014
# WISC SUBTESTS Significant at .01 level	8			1	7			10		

Table 14

Correlations and p Values for Items in the MAP Non Verbal Index

Correlations significant at .01 or better

	SEQUENCING			BLOCK TAPPING			OBJECT MEMORY			PUZZLES			FIGURE GROUND		
	W WITHDRAW P	.12	p=.01	PERFORM IQ	.23	p=.001	WJ LANG	.09	p=.038	FULL IQ	.33	p=.001	WJ LANG	.30	p=.001
				WJ LANG	.19	p=.001	W PEER P	.09	p=.057	PERFORM IQ	.31	p=.001	FULL IQ	.26	p=.001
				T MOTIV	.18	p=.001	BRUIN TOTAL	.09	p=.051	VERBAL IQ	.29	p=.001	PERFORM IQ	.25	p=.001
				T ATTN SPAN	.17	p=.001	BEERY	-.11	p=.018	WJ READING	.27	p=.001	VERBAL IQ	.23	p=.001
				VERBAL IQ	.17	p=.001	T PROB SOLV	-.11	p=.030	WJ LANG	.24	p=.001	BRUIN GROSS	.18	p=.001
				WJ READ	.17	p=.001				WJ MATH	.23	p=.001	BRUIN TOTAL	.15	p=.003
				WJ MATH	.16	p=.002				BRUIN GROSS	.20	p=.001	WJ MATH	.15	P=.003
				T PROB SOLV	.16	p=.002				T READING	.20	p=.001	BEERY	.14	p=.006
				DAP #9	.15	p=.003				BEERY	.19	p=.001	PHYS T	.13	p=.015
				M COG T	.15	p=.004				GOODENOUGH	.19	p=.001	W DISTRAC P	.13	p=.007
				T HANDWRT	.14	p=.006				T VOCAB	.18	p=.001			
				T BAL/COORD	.14	p=.008				BRUIN TOTAL	.18	p=.001			
				T MEMORY	.14	p=.006				DAP #3	.18	p=.005			
				M PHYS T	.13	p=.001				T ATTN SPAN	.16	p=.002			
				T VERBAL	.13	p=.008				T DEXTER	.16	p=.002			
				T PHYS ABIL	.13	p=.01				T PROB SOLV	.16	p=.003			
				T LANG	.13	p=.008				M COG T	.15	p=.004			
				T READ	.12	p=.014				T OVERALL	.13	p=.010			
				T PE	.12	p=.01				W PEER P	.13	p=.008			
				BEERY	.12	p=.012				W DISTRAC T	.12	p=.017			
				W DISTRAC P	-.16	p=.002				W TOTAL P	.12	p=.015			
										T LANGUAGE	.12	p=.014			
# WISC SUBTESTS Significant at .01 level	2			9			0			10			10		

Table 15

Correlations and p Vaules for Items in the MAP Complex Tasks Index

Correlations significant at .01 or better

BLOCK DESIGNS		
PERFORM IQ	.28	p=.001
FULL IQ	.27	p=.001
WJ MATH	.25	p=.001
BRUIN TOTAL	.24	p=.001
VERBAL IQ	.22	p=.001
H PHYS T	.21	p=.001
GOODENOUGH	.20	p=.001
T VOCAB	.19	p=.001
DAP #3	.18	p=.005
WJ READ	.18	p=.001
DAP #10	.17	p=.009
T PROB SOLV	.17	p=.001
WJ LANG	.16	p=.002
T ATTN SPAN	.16	p=.002
BRUIN FINE	.15	p=.002
BRUIN GROSS	.14	p=.005
H COG T	.14	p=.002
W IMMAT P	.14	p=.006
T OVERALL	.14	p=.006
BEERY	.13	p=.009
H SOC T	.13	p=.011
W TOT P	.13	p=.010
T READING	.13	p=.010
T VERBAL	.12	p=.013
T BAL/COORD	.12	p=.014
T DEXTER	.12	p=.014

DRAW A PERSON		
FULL IQ	.23	p=.001
PERFORM IQ	.22	p=.001
VERBAL IQ	.21	p=.001
WJ MATH	.19	p=.001
H PHYS T	.18	p=.001
DAP #3	.17	p=.009
BRUIN TOTAL	.16	p=.002
BRUIN FINE	.13	p=.008
WJ READ	.13	p=.008
W PEER P	.13	p=.007
T READING	.13	p=.012
GOODENOUGH	.12	p=.017
BRUIN GROSS	.11	p=.022
WJ LANG	.11	p=.019

IMITATE POSTURES		
DAP #3	.18	p=.007
DAP #5	.15	p=.015
BRUIN GROSS	.14	p=.005
W IMMAT P	.13	p=.009
T BAL/COORD	.13	p=.012

MAZE		
PERFORM IQ	.28	p=.001
FULL IQ	.25	p=.001
WJ READ	.21	p=.001
DAP #5	.20	p=.002
T VOCAB	.20	p=.001
BEERY	.19	p=.001
VERBAL IQ	.19	p=.001
WJ MATH	.19	p=.001
WJ LANG	.19	p=.001
H COG T	.19	p=.001
T READING	.19	p=.001
BRUIN TOTAL	.17	p=.001
DAP #3	.17	p=.009
T OVERALL	.17	p=.001
T VERBAL	.17	p=.001
DAP #10	.16	p=.014
T LANGUAGE	.16	p=.001
T WRIT EXP	.16	p=.002
BRUIN FINE	.15	p=.004
GOODENOUGH	.14	p=.005
W WITHDRAW T	.14	p=.008
W WITHDRAW P	.12	p=.012

	BLOCK DESIGNS	DRAW A PERSON	IMITATE POSTURES	MAZE
# WISC SUBTESTS Significant at .01 level	10	7	1	8

In the following section, the findings and results are discussed for each of the twenty seven items of the MAP, listed in alphabetical order. The abbreviation following the item name refers to the index (or subtest) from which the item comes.

Articulation (CI):

Articulation correlated with 26 measures and all 10 subtests from the WISC-R. The highest correlations were seen between Articulation and the IQ and achievement measures (r =.34 to .24). Next the Bruininks Total score (r = .21) and Bruininks Gross motor score (r = .20) were highly significant. Seven of the teacher checklist items were highly significant including: verbal expression, language arts, handwriting, written expression, overall rating, math, and vocabulary (r = .21 to .13).

Other significant correlations included the Beery and the Goodenough (r = .20 and .18), and five of the Harter subtests: Physical teacher, Social teacher, Cognitive teacher, Physical child, and Social child (r = .19 to .13). The items on the Walker that were significant included: Withdrawal teacher, Total score parent, Acting out parent, and Distractability parent (r = .13 to .12).

Articulation had the highest correlations with cognition and achievement tests, however, it also appeared to be closely related to motor skills. Five of the seven teacher ratings which correlated at a significant level were related to language, and it seems likely that either articulation disorders are associated by teachers with language problems, or in fact the two types of problems coexist. This item had a larger number of high correlations with the Harter than almost any other MAP item, suggesting that children with articulation disorders tend to have problems with self image, more so than children who did poorly on other MAP items. Articulation is also one of the few items which correlated to the Beery and Goodenough at a significant level, suggesting a relationship between oral motor and fine motor.

Figure 1

Abbreviations for Selected Criterion Measures in Appendix C

Teachers Observations of Language Arts	T Language
Teachers Observations of Reading	T Reading
Teachers Observations of Math	T Math
Teachers Observations of Physical Education	T PE
Teachers Observations of Overall Academic Performance	T Overall
Teachers Observations of Verbal Expression	T Verbal
Teachers Observations of Handwriting	T Hand Wrtg
Teachers Observations of Written Expression	T Wrt Exp
Teachers Observations of Vocabulary	T Vocab
Teachers Observations of Problem Solving	T Prob Solv
Teachers Observations of Memory	T Memory
Teachers Observations of Following Directions	T Fol Dir
Teachers Observations of Motivation	T Motiv
Teachers Observations of Balance/Coordination	T Bal/Coord
Teachers Observations of Attention Span	T Attn Span
Teachers Observations of Physical Activities and Sports	T Phys Act
Teachers Observations of Dexterity	T Dexter
Harter Physical Ability Teacher	H Phys T
Harter Physical Ability Child	H Phys C
Harter Peer Relationships Teacher	H Peer T
Harter Peer Relationships Child	H Peer C
Harter Cognitive Abilities Teacher	H Cog T
Harter Cognitive Abilities Child	H Cog C
Harter Social Abilities Teacher	H Soc T
Harter Social Abilities Child	H Soc C
Harter Appearance Teacher	H App T
Harter Appearance Child	H App C
Harter Conduct Teacher	H Cond T
Harter Conduct Child	H Cond C
Harter Maternal Attachment Child	H Mat Att C
Walker Peer Relationships Teacher	W Peer T
Walker Peer Relationships Parent	W Peer P
Walker Immaturity Teacher	W Immature T
Walker Immaturity Parent	W Immature P
Walker Acting Out Teacher	W Act Out T
Walker Acting Out Parent	W Act Out P
Walker Distractability Teacher	W Distrac T
Walker Distractability Parent	W Distrac P
Walker Withdrawal Teacher	W Withdraw T
Walker Withdrawal Parent	W Withdraw P
Walker Total Score Teacher	W Total T
Walker Total Score Parent	W Total P

The Bruininks-Oseretsky is abbreviated as Bruin, the Woodcock-Johnson is abbreviated as WJ or Woodcock, the Goodenough-Harris is abbreviated as Goodenough, and Intelligence Testing is abbreviated as WISC-R on tables 11-15.

Block Designs (CTI):

Block Designs correlated at a highly significant level with all 10 WISC-R subtests, and 26 criterion measures, the highest of which were Performance IQ, Full Scale IQ, Woodcock Math, and the Bruininks motor scale (r = .28 to .24). Highly significant also were relationships with Verbal IQ, the Harter Physical Teacher score, and the Goodenough (r = .22 to .20). Eight of the teacher checklist items are significant including: vocabulary, problem solving, attention span, overall rating, reading, verbal expression, balance/coordination, and dexterity (.19 to .12). The Woodcock reading and language scores correlated at a highly significant level also (r = .18 and .16). Both the Bruininks gross motor and fine motor scores correlated at a highly significant level at .15 and .14. Two of the Harter items correlated at a highly significant level: Cognitive teacher, and Social teacher. Two of the Walker items correlated significantly: Immaturity parent, and Total parent. In addition two of the twelve qualitative aspects of figure drawing correlated significantly: #3,draws a monster or dinosaur, etc., and # 10, unenclosed (r = .18 and .17).

These findings suggest that Block Designs is an antecedent of performance and math skills, more than verbal abilities. Interestingly the Performance IQ correlation is larger than the Verbal IQ measure, the teacher checklist items were not exclusively verbal, the Harter score was physical, and the Bruininks correlations were relatively large. Block Designs was included in the CTI because it seemed highly visual, as well as motoric. It is interesting to note that it did correlate significantly with the Beery and the Goodenough, but not as significantly as it did with performance measures.

Block Tapping (NVI):

Block Tapping correlated with 21 measures, and 9 WISC-R subtests. Although the total number of correlations was high, the correlational values were somewhat lower than some of the correlations with other MAP items, yet still highly significant. Performance IQ was the most highly correlated

(.23), with Woodcock Language next (.19). Ten of the teacher ratings correlated at a highly significant level including: motivation, attention span, problem solving (r = .17 to .16),handwriting, balance/coordination, memory, verbal expression, physical abilities, language, reading, and physical education (r = .14 to .12). Verbal IQ, and Woodcock Reading and Math were highly significant at .17 and .16. Two of the Harter items were significant: Cognitive teacher, and Physical teacher. The Beery was significant but not highly (r = .12).

This item was hypothesized to measure memory and sequencing, but interestingly, seems to be antecedent to performance and language. It is also interesting that the teachers' observations with the largest correlations are not academic, but rather behavioral.

Cage (CI):

Cage correlated with the more criterion measures, 31, than any other MAP item. It also correlated with six WISC-R subtests at the .01 or better level, and the rest of the WISC-R subtests at the .05 level. It is one of the only items for which the highest correlations were not the IQ measures. The highest correlations were with the teacher observations of math and attention span (r = .26) and with the Walker Peer Relationships parent (r = .24). Also highly significant were the Bruininks Total score, Bruininks Fine motor score, and Bruininks Gross motor score (r = .23, .20, and .17 respectively). Woodcock Reading and Math, and the Goodenough were highly significant (r = .21 to .20). Fourteen teacher observations were significant, the two mentioned above, and: reading, overall rating, balance/coordination, hand writing, written expression, language, follows directions, problem solving, dexterity, vocabulary, motivation, and memory. Performance IQ, Full scale IQ, and Verbal IQ were highly significant but not near the top of the list (r = .18, .17, and .14). Three additional Walker subtests were highly significant: Total parent, Immaturity parent, and Distractibility parent. Woodcock Language and Beery appear low on the list (r = .16 and .13).

These findings differ from most of the MAP items, which tend to correlate most highly with IQ and achievement. This simple fine motor task does indeed seem to be a precursor of motor skills, but interesting also seems to tie into teacher perceptions of the child in unexpected ways such as attention span. It may be that the skills needed to perform this task in preschool years involve attention, rather than more cognitive abilities.

Digit Repetition (VI):

This item correlated with all 10 WISC-R subtests at a highly significant level, and with 28 other measures. The Woodcock Math, Reading and Language correlations were among the largest as expected, (r = .32 and .30) as were the Verbal IQ and Full Scale IQ scores (r = .31 and .29). Surprisingly, the Bruininks total is highly significant (r = .21) with the other Bruininks scores also significant (Fine motor = .14, and Gross motor = .12). This item appeared to be quite predictive of a variety of teacher observations including: language arts, reading, written expression , math, overall rating, verbal expression, memory, motivation, vocabulary, problem solving, and attention span (r = .21 to .12). Performance IQ was highly significant but lower than Verbal IQ (r = .20). Two of the Harter subtests were correlated at a highly significant level, Cognitive teacher, and Cognitive child (r = .20 and .12). Four of the Walker subtests correlated significantly, Acting Out parent, Total parent, Acting Out teacher, and Peer Relationships parent (r = .15 to .12). The Beery and Goodenough both correlated at the same level (r = .13).

The primary correlates, IQ and achievement, are logical since Digit repetition seems to require similar skills such as memory and verbalization to the skills required on IQ and achievement testing. Digit repetition (or the abilities it measures in the preschool years) may be an antecedent for what is measured and later called "intelligence". The level of correlation with the motor measures and with Performance IQ, about .20, was somewhat unexpected since this item does not appear to have an motoric components. This item was predictive of 11 teacher observations, making it one of the itmes with the highest number of significant correlations with teacher observations later.

Draw A Person (CTI):

This item correlated with 15 criterion measures, and 7 of the WISC-R subtests. The highest correlations were with the three IQ measures (r = .23 to .21), and with the Woodcock Math score (r = . 19). There was a highly significant correlation between Draw A Person and the Harter Physical teacher subtest (r = .18) as well the qualitative draw a person item #3, draws monster, dinosaur etc. (r = .17). The three Bruininks scores correlated significantly, Total score at .16, Fine motor at .13 and Gross motor at .11 (The latter score is significant at the .05 level only.). The Goodenough score was correlated significantly at .12, as was the Woodcock Language at .11. In addition, the Teachers' observation of reading (r = .13) and the Walker Peer Relationships parent (r = .13) were significant.

Of interest in this item is the relatively low, though significant correlation between the MAP Draw a person, and the Goodenough Draw a person. Interestingly, Draw a person in the early years correlates more highly with WISC-R scores four years later than with figure drawing four years later. This figure drawing ability also appears to be related to later motor performance, which is logical since it has a large motor component.

Figure Ground (NVI):

Figure Ground correlated with 10 measures, and all 10 WISC-R subtests. The correlations with Woodcock Language, and all three IQ measures were particularly high (r = .30, .26,.25,and.23). Not as high, but also significant were correlations with the Bruininks Gross motor and Total score (r = .18 and .15). The Woodcock Math, and Beery were significant at about the same level (r = .15 and .14). The Harter Physical Teacher score and Walker Distractability parent score were also significant (r = .13).

This item appears to be an early precursor of intelligence as measured by the WISC-R, and of language in particular as measured by the Woodcock. This is quite an interesting finding considering the apparent visual perceptual nature of the task. It appears likely that the skills necessary to complete visual perceptual tasks in the preschool years (such as the development of mental imagery) may later be the abilities used to complete tasks measuring general intelligence.

Finger Localization (FI):

This item only correlated with four criterion measures, and three of the WISC-R subtests. The highest correlation was seen between Finger localization and the Draw a person qualitative item #1, body parts not attached (r = .24). In addition Performance IQ and Full Scale IQ were significantly correlated (r = .15 and .12). The Harter Social child scale also correlated at a significant level.

This is one of the items on which longitudinal research has been reported previously. Although the findings that Finger localization is an early antecedent of intelligence was supported in part by the data in this study, it does not seem to be as good a predictor of WISC-R scores as many of the other items on the MAP. In particular the "better" performance of Stereognosis, which also measures tactile skills, but requires a higher degree of integrative skill, was noted.

Follow Directions (VI):

This was the only MAP item which did not correlate with any of the criterion measures, and with only one of the WISC-R subtests, which could have been a function of chance.

This is an extremely interesting finding because this item is frequently administered in the preschool years, and is generally felt to be quite indicative of future abilities (see Review of the Literature, Review of Language Abilities). The Follow Directions item on the MAP does not appear to be administered in an unusual manner, or poorly constructed psychometrically. Thus the only logical conclusion is that whatever this item measures in the preschool years, it is not a correlate of what any of the criterion measures tested.

General Information (VI):

This item correlated with 15 criterion measures, and eight of the WISC-R subtests. The highest correlations were with Verbal IQ, Woodcock Reading, and Full Scale IQ ($r = .20$ to .19). The Bruininks Total score and Gross Motor score were correlated at .17 and .13 respectively. The Woodcock Language and Math scores were highly significant ($r = .17$ and .15), although not as high as the Woodcock Reading score. Five of the Walker scores correlated, more than correlated with most of the MAP items: Walker Withdrawal parent, Walker Acting Out teacher, Walker Total teacher, Walker Immaturity teacher, Walker Total Parent ($r = .17$ to .12). Performance IQ correlated at a significant level, but lower than Verbal IQ at .14. Two of the Harter subtests correlated significantly, the Cognitive teacher, and the Social teacher.

These findings suggest that General Information as expected correlated highly with Verbal IQ and reading. It is interesting to note the number of Walker and Harter subtests that are significant. This suggests that there are behavioral correlates in primary school years for what is measured by General Information in the preschool years.

Hand-Nose (FI):

Hand-Nose, known as finger-nose by neurologists, correlated with 14 criterion measures, and 7 WISC-R subtests. The highest correlations were with Verbal IQ and Full Scale IQ ($r = .21$ and .20). The Bruininks Gross motor and Total scores were also highly significant ($r =.20$ and .18), although the Bruininks Fine motor score was not. The Woodcock Math, Language, and Reading scores were highly significant although the correlations were not as high as those mentioned above (r = .17,.16,and.14). Performance IQ did not correlate as well as Verbal IQ, although it was still significant (r = .16). Four of the teachers' observations correlated highly: memory, balance/coordination, physical activity, and dexterity (r = .15 to .13). The Harter Peer Relationships teacher (r = .14) and the Walker Acting Out parent were also significant (r = .14 and .12).

These findings are interesting, particularly the high correlation between Hand-Nose, a standard neurological item, and the IQ measures. As with several of the other neurological items mentioned below (Stereognosis, Romberg, and Stepping), the high correlation may suggest that there are neurological foundations to intelligence (as measured by the WISC-R). The motoric relationships are not surprising, and suggest that this task is also a precursor of later motor functioning. It is noteworthy that three of the four significant teachers' observations involved motor performance.

Imitation of Postures (CTI):

Only five of the criterion measures correlated with this item, and 1 of the WISC-R subtests. The highest correlations were with two of the Draw a person qualitative checklist items: #3, Draws monster, dinosaur, etc. and #5, not identifiable as a human figure (r =.18 and .15). The Bruininks correlated at a significant level at .14. In addition, the Walker Immaturity parent subtest, and the teachers' observations of balance and coordination were significant (r =.13).

This item appears to be the least predictive of the criterion measures of any item in the CTI. This may have to do with the nature of the administration and scoring of this item, which is a little more subjective than the other MAP items, or perhaps with the content of what is measured. The significant correlations with the Bruininks, and the teacher observation of balance and coordination are logical.

Kneel-Stand (FI):

This item was one of the least correlative items, correlating with seven criterion measures, and six of the WISC-R subtests. The highest correlations were with the Bruininks Gross motor score (r = .19), and with Performance IQ and Full Scale IQ (r = .15). The Bruininks Total score and Verbal IQ were also significant (r = 13 and .12). The Harter Social child, and teachers' observations of dexterity were correlated at a significant level (r = .14, and .13).

This item was designed to measure rotation (or the ability of the child to cross the midline of their body), and thus since it is strictly a motor measure, it is not surprising that it does not correlate with most of the criterion measures, which are not motoric. The high correlation with the Bruininks scores is logical, since the Bruininks is a motor test. It is interesting that the Performance and Full Scale IQ correlate as highly as they do in this task, since the task does not appear on the surface to have much in common with intelligence measures.

Maze (CTI):

Maze correlated at a highly significant level with 22 of the criterion measures, and eight of the WISC-R subtests. It appeared to be second only to Block Designs in this index for number of significant correlations. The highest correlations were with Performance IQ, Full Scale IQ, and Woodcock Reading (r = .28, .25, and .21). Three of the qualitative aspects of Draw a person were significant: #5, not identifiable as a human figure, #3, draws monster, dinosaur etc., and #10, unenclosed (r = .20, .17, and .16). Six of the teachers' observations were significantly correlated: vocabulary, reading, overall rating, verbal expression, language, and written expression (r = .20 to .16). The Beery, Verbal IQ score, Harter Cognitive teacher, Woodcock Math and Woodcock Language were similarly correlated at .19. The Bruininks Total and Fine motor scores were significant (r = .17 and .15). In addition the Goodenough was highly significant at .14, as were the Walker Withdrawal teacher and Walker Withdrawal parent subtests (r = .14 and .13).

This interesting motor planning item involved a large component of cognitive activity. It strongly correlates with IQ, particularly Performance IQ as measured by the WISC-R. It was put into the CTI since it seemed to combine motor and cognitive skills, however on the basis of this information it seems likely that it could have been placed in the NVI. The correlations with motoric criterion measures were relatively low compared to the correlations with cognitive criterion measures.

Object Memory (NVI):

This item appeared to be the poorest one in this index according to this correlative information. It did not correlate at a significant level with any of the criterion measures, or with any of the WISC-R subtests.

This finding was quite unexpected as Object Memory is a familiar item, administered frequently in testing for young children. From this data it does not appear to have a common variance with, or degree of overlap with any of the criterion measures.

Puzzles (NVI):

Puzzles correlated significantly with 22 of the criterion measures, and all 10 WISC-R subtests. In addition to Block Tapping, this appeared to be the strongest item in this index, and one of the best items in the test. Extremely high correlations were seen with all three IQ measures (r = .33 to .29), and with all three achievement measures of the Woodcock (r = .27 to .23). The Bruininks Gross motor and Total score correlated highly at .20 and .18 respectively. Seven of the teachers' observations correlated significantly: reading, vocabulary, attention span, dexterity, problem solving, overall rating and language arts (r = .20 to .12). The Beery and Goodenough were correlated similarly at .19. One of the Harter subtests, Cognitive teacher, and three of the Walker subtests, Peer Relationships parent, Distractability teacher, and Total parent, correlated at a significant level (r = .15 to .13).

This task appears to be directly related to IQ and achievement as measured by the WISC-R and Woodcock. Although the task is visual perceptual in nature, and is not a timed task, it appears to be an early precursor of the skills later measured in intelligence testing. The pattern with this item is similar to that of Figure Ground, and it is possible with both items that this is related to the development of mental imagery. With the exception of Articulation, this item has the highest correlations to the three IQ measures.

Romberg (FI):

Romberg correlated with 21 of the criterion measures, and seven of the WISC-R subtests. Interestingly, the highest correlation was with the Harter Physical teacher subtest (r = .28), and several other Harter Subtests were also highly significant: Maternal Attachment (r = .20); Cognitive teacher (r = .15), and Social teacher (r = .12). The Bruininks Total score correlated most highly of the standardized test scores (r = .22), and the Bruininks Gross motor correlation was highly significant at .19, as was the Bruininks Fine motor correlation at .14. Two of the Draw a person qualitative items were highly significant, #12, Body parts definitely out of proportion, and # 9, Bizarre, disturbing quality (r = .21 and .18 respectively). The three IQ measures were highly significant at .19 (Full Scale), .18 (Verbal Score), and .17 (Performance Score).

Five of the teachers' observations were significant: language, math, overall rating, follow directions, and reading (r = .18 to .12). Two Walker subtests were highly significant, Peer Relationships teacher, and Total teacher (r = . .16 and .15). Two of the Woodcock tests were highly significant, although the correlations were lower than with many of the other MAP items, Math (r = .15), and Language (r = .14).

This item is an unusual MAP item, because its correlative focus is with motor measures including standardized motor tests, and qualitative aspects of physical functioning. It is somewhat surprising that the IQ and achievement measures correlate as highly as they do with this item, since Romberg is clearly a neurological test item. This data lends further support to the theory that there are neuromaturational substrates in the preschool years that provide some of the foundations for later academic performance.

Sentence Repetition (VI):

This item correlated at a significant level with twelve of the criterion measures, and seven of the WISC-R subtests. The highest correlations were with the Woodcock Reading and Math subtests (r = .23 and .21), and the the Verbal IQ and Full Scale IQ (r = .22 and .20). The Woodcock Language and Performance IQ were also highly significant at .17, and .14. Interestingly, there appears to be a relationship between Sentence repetition and motor abilities since the Bruininks Total score, Gross motor and Fine motor scores were all highly correlated (r = .20, .17, and .12). Two of the teachers' observations were highly significant: reading, and language (r = .14 and .13). In addition the Draw a person qualitative item, # 12, Body parts definitely out of proportion, was also significant (r = .16).

As expected, this item correlated well with general measures of intelligence and achievement. Of more surprise were the relatively high correlations with the motor measures, suggesting that performance of this item in the early years may be a precursor to more than simply verbal abilities.

Sequencing (NVI):

This item correlated with only one criterion measure, the Walker Withdrawal parent subtest (r = .12). Two of the WISC-R subtests correlations were high enough to reach significance.

It is possible that these results could occur on the basis of chance alone. It is interesting that an item such as Sequencing which is documented to be a problem in many school age learning disabled children (see review of the literature, Chapter II), is apparently not predictive of any of the criterion measures employed in this study. This may be a function of the MAP test item administration and scoring instructions, or may actually indicate that sequencing is not an antecedent of the domains measured in this study.

Stamp (CI,FI):

Stamp correlated at a highly significant level with eleven of the criterion measures, and all ten WISC-R subtests. The highest correlations were with the IQ measures: Verbal IQ and Full Scale IQ (r = .24), and Performance IQ (r = .19). High correlations were also observed with the Bruininks Total score, Gross motor, and Fine motor scores (r = .21, .20, and .13). The Woodcock Math, Reading, and Language correlations were also highly significant (r = .18 to .17). Two of the Walker subtests were significant, Withdraw teacher, and Immaturity teacher (r = .12).

These were surprising findings for this item which seems to be assessing motor coordination. The highest correlations were with IQ, and the correlations with achievement are also surprisingly high. This suggests that the ability to perform rapid alternating movement patterns such as Stamp, may be a precursor or foundation for what is later measured by the WISC-R and Woodcock, and commonly called, intelligence and achievement.

Stepping (FI):

This item correlated at a highly significant level with 13 of the criterion measures, and five of the WISC-R subtests. The highest correlation was with the Harter Physical child subtest (r = .20), and five other Harter subtests also correlated at a high level: Cognitive teacher, Appearance child, Cognitive child, Social child, and Physical teacher (r = .16 to .13). The IQ measures were highly significant at .18,.17 and .16. Woodcock Math, and Reading were significantly correlated at .17 and .13 respectively. One of the qualitative aspects of Draw a person was correlated, #11, vague overall shape, but hard to distinguish individual parts (r = .16). The only motor score that was correlated was the Bruininks Total score (r = .14). In general, although the correlations are highly significant, they are lower than for many of the other MAP items.

Interestingly, this item is a standard neurological item, administered to adults to determine possible cerebellar dysfunction. Yet it correlated with IQ, and with self perception. The motor score was among the lowest of the significant correlations. Stepping appears to be a motor task, but possibly the neurological demands made on the child have something in common with the demands later made during standard intelligence testing tasks.

Stereognosis (FI):

This integrative tactile task correlated with 14 of the criterion measures, and all 10 WISC-R subtests. Remarkably, in a task which appears to be sensory motor in nature, the highest correlations were all three IQ measures (r = .26,.24,.23). In addition quite high correlations with Woodcock Math, and Language were noted (r = .22 and .18). Two of the Walker subtests were correlated at a highly significant level, Distractability parent, and Total parent (r = .18 and .15). The Bruininks Total score, and Gross motor scores were correlated highly (r = .17 and .12). One of the Draw a person qualitative items was correlated at a significant level, # 9, Bizarre, disturbing quality (r = .17). Three of the Harter subtests were correlated at a significant level, Peer Relationships teacher, Peer Relationships child, and Cognitive teacher (r = .16, .15, and .13). The teachers' observation of verbal expression was also correlated at a significant level (r = .12).

This tactile item assesses a behavioral domain which has been hypothesized to be one of several sensory systems which forms the foundations for later academic learning. It is quite interesting to note the high correlations with IQ and math and language. These data are suggestive that this task, although it appears to be sensory, may in fact be one of the precursors for abilities later measured by intelligence testing.

Supine Flexion (FI):

Supine Flexion correlated at a highly significant level with fourteen of the criterion measures, and three of the WISC-R subtests. The highest correlation unexpectedly was with the Harter Conduct child subtest (r = .22), and two other Harter subtests also were highly significant: the Physical child, and the Physical teacher (r = .15). As expected the Bruininks correlated quite highly: Total score r = .19; Gross motor score r = .16, and Fine motor score r = .14. This was one of the only items that correlated at a significant level with either the Beery or the Goodenough (r = .15). The three Woodcock tests correlated at a highly significant level, but not as highly as with most of the other items, (r = .12 to .14) Three of the teachers' observations correlated at a highly significant level: balance/coordination, physical activity, and dexterity (r = .13).

This item, like Romberg, seems unlike most of the other MAP items, in that although it correlates with a high number of criterion measures, they are primarily physical in nature. The Bruininks motor test, two of the three Harter subtests, and all three significant teachers' observations are physical. This is quite logical since Supine flexion measures physical strength, and reflex integration. It seems likely that this item will contribute to the MAP total score something different than what is contributed by most of the other items.

Tongue Movement (CI):

This item correlated at a highly significant level with five of the criterion measures, and three of the WISC-R subtests. The two highest correlations were with two of the Draw a Person qualitative items, #12, Body parts definitely out of proportion, and # 5, Not identifiable as a human figure (r = .19 and .17). Performance IQ, the Goodenough, and teachers' observations of handwriting were all significant at the same level (r = .12).

This does not appear to be as good an item as most of the MAP items, in terms of an overlap in variance with the criterion measures in this study. Although it is unlikely that all five of these high correlations could have happened by chance, there is not a clear pattern in the measures which correlated. This item does not appear to be a precursor of a particular type of later functioning.

Tower (CI):

Tower correlated with thirteen of the criterion measures, and nine of the WISC-R subtests. Five of the teachers' observations were significant at a high level: language (.21), balance/coordination (.19), written expression (.15), memory (.13), and physical activity (.12). The Bruininks Total score was highly correlated, as was the Harter Physical teacher subtest (r = .20). The three IQ measures were highly correlated at .17 to .15, for Full Scale, Verbal, and Performance. The Woodcock measures were also highly significant: Reading (.15), Language (.12), and Math (.11).

Not unexpectedly, there is a preponderance of correlations with physical measures. However, it is interesting to note that this item also correlates significantly with IQ and with achievement. Perhaps this fine motor task is necessary for adequate overall functioning in the preschool age group, and children who do poorly at this task in the early years are likely to experience problems in several domains later.

Vertical Writing (FI,CI):

This item correlated only with the Bruininks Fine motor score (r = .12), and with one WISC-R subtest. Although this correlation could have happened by chance, it is a logical one since this task is fine motor. However, based on correlative information, this task in the preschool years does not seem to be intrinsically valuable.

Walks Line (FI,CI):

Walks Line correlated at a highly significant level with five of the criterion measures, and one WISC-R subtest. The highest correlation was with the Beery, the Walker Acting Out parent and Walker Total parent subtests (r = .13). In addition, the Bruininks Gross motor score and the Woodcock Math score were correlated at a highly significant level (r = .12).

This does not appear to be one of the better MAP items from these analyses. Although highly significant, the correlations are relatively low. This item does not appear to be highly indicative of future abilities, as measured by the criterion measures in this study.

Research Question #2

How well does the MAP discriminate between problem and no problem students? That is, can the MAP distinguish between the children who perform in the low range and the children representing the rest of the distribution?

The discussion of this Research Question is subdivided into eight parts, corresponding to the eight problem/no problem categories operationally defined below:

1. Retained In School

The children in the Retained In School group were described by teachers, parents, or in the school records as having "flunked" or been held back for one or more grades in school.

2. Failed Teachers' Observations

The children in the Failed Teachers' Observations group were rated "below average" by their teachers on 50% or more of the following school subjects and behavioral characteristics: language arts, reading, math, physical education, overall academic performance, verbal expression, handwriting, written expression, vocabulary, problem solving, memory, following directions, motivation, balance/coordination, attention span, physical activities/sports, and dexterity. (As noted in Chapter III, each of these checklist items were rated by teachers as below average, average, or above average.)

3. Received Special Services

The children in the Received Special Services group were described by the teachers, parents, or in the school records as receiving one of the following: Speech therapy, Occupational therapy, Physical therapy, Tutoring, Psychological counseling, Remedial reading, or Adaptive physical education.

4. In A Special Class

The children in the In A Special Class group were noted by teacher, parent, or in the school records as being in one of the following self contained classrooms: mentally retarded, emotionally disturbed, or learning disabled.

5. Failed Report Card Language

The children in the Failed Report Card Language group were noted in the school records as having received a grade representative of below average status in any grade level in Language Arts.

6. Failed Report Card Reading

The children in the Failed Report Card Reading groups were noted in the school records as having received a grade representative of below average status in any grade level in Reading.

7. Failed Report Card Math

The children in the Failed Report Card Math groups were noted in the school records as having received a grade representative of below average status in any grade level in Reading.

8. Failed Report Card Physical Education

The children in the Failed Report Card Physical Education groups were noted in the school records as having received a grade representative of below average status in any grade level in Physical Education.

The number of children in each of the problem/no problem categories are described in Table 16. As can be seen, in the Retained In School category, 14% of the sample were in the problem group; in the Failed Teachers' Observations group, 8% of the sample was in the problem group; forty percent of the sample Received Special Services; Special Classes were needed by 19% of the sample; the failed group for each of the Report Card categories was as follows: language - 7%, reading - 20%, math - 8%, and physical education 2%. Therefore it can be seen that depending upon the way the problem category was defined a range of 2 to 40 percent of the sample were in the problem group.

To a certain extent, Research Question #2 represents a shift in perspective from that of Research Question #1. As stated previously the criterion measures used in Research Question #1, were considered to be correlates or direct measures of aptitude. A high correlation between the MAP and these measures would indicate that the MAP, like the criterion measures, will accurately rank order a wide range of children relative to their aptitude. Research question #2 was meant to determine whether students from the lower range of the aptitude distribution could be distinguished from the students from the rest of the distribution using the MAP. The problem groups as described above represented the low aptitude students; the no problem groups represented the remainder of the students.

To answer Research Question #2, the following analyses were undertaken, and findings and conclusions are described below for each of the eight problem/no problem categories:

a) Determination of whether the problem/no problem groups represented different aptitude populations by conducting two tailed, t-tests for independent groups using the following dependent measures: chronological age, Berry Visual Motor, WISC-R Full Scale IQ, Goodenough Drawing, Woodcock Reading, Woodcock Math, and Woodcock Language. Table 16 reports the results of these analyses for each of the eight problem vs. no problem groups.

b) Determination of whether the MAP Total Score, and MAP Indices (subtests) can discriminate between the problem/no problem groups by conducting two tailed, t-tests for independent groups using the MAP Total score and the MAP Indices as dependent measures for each of the eight problem categories. Table 17 provides the results for these analyses, including t values, p values, and mean scores for groups.

Table 16

Number of Children Delineated in the Problem vs. No Problem Groups

Retained In School	N	%
N: Not Retained	291	86
P: Retained	47	14

Teacher Observations	N	%
Passed	311	92
Failed	27	8

Special Services	N	%
Not Receiving	201	60
Receiving	137	40

Special Class	N	%
Not Needed	273	81
Needed	65	19

REPORT CARDS

Language	N	%
N: Passed	313	93
P: Failed	25	7

Reading	N	%
Passed	318	94
Failed	20	6

Math	N	%
Passed	312	92
Failed	26	8

Physical Education	N	%
Passed	332	98
Failed	6	2

N: No Problem
P: Problem

Table 17
Determination of Whether Problem/No Problem Groups Represent Different Populations

N = No Problem Group
P = Problem Group

Dependent Measure	Retained In School		Teacher Observations		Special Services		Special Class	
	Means	SD	Means	SD	Means	SD	Means	SD
Chronological Age	N 8.51 P 8.47	.83 .88	N 8.46 P 8.92	.83 .71	N 8.44 P 8.59	.83 .84	N 8.45 P 8.70	.84 .77
	t=.30	p=.767	t=2.81	p=.005	t=1.62	p=.106	t=2.19	p=.029
Beery Visual Motor	N 9.54 P 8.36	3.06 3.00	N 9.61 P 6.59	3.01 2.39	N 10.11 P 8.28	3.04 2.79	N 9.74 R 7.80	3.05 2.66
	t=2.45	p=.015	t=5.08	p=.000	t=5.61	p=.000	t=4.70	p=.000
WISC-R Full Scale IQ	N 116.74 P 105.04	15.51 17.28	N 116.39 P 100.44	14.92 22.97	N 119.01 P 109.40	14.30 17.27	N 117.24 P 106.17	14.59 19.65
	t=4.72	p=.000	t=5.07	p=.000	t=5.57	p=.000	t=5.12	p=.000
Goodenough Drawing	N 99.75 P 91.47	16.27 16.48	N 99.56 P 87.48	16.18 16.64	N 100.74 P 95.45	17.11 15.14	N 99.78 P 93.63	16.61 15.32
	t=3.23	p=.001	t=3.71	p=.000	t=2.92	p=.004	t=2.72	p=.007
Woodcock Reading	N 106.81 P 95.72	12.72 14.51	N 106.55 P 90.59	12.70 14.59	N 108.57 P 100.44	12.21 13.92	N 108.00 P 93.82	11.91 13.95
	t=5.43	p=.000	t=6.20	p=.000	t=5.67	p=.000	t=8.34	p=.000
Woodcock Math	N 107.26 P 94.34	13.54 15.15	N 106.57 P 92.74	13.67 17.32	N 108.58 P 100.89	12.94 15.38	N 107.69 P 96.12	12.97 16.64
	t=5.97	p=.000	t=4.93	p=.000	t=4.97	p=.000	t=6.10	p=.000
Woodcock Language	N 108.92 P 95.55	13.95 14.12	N 108.36 P 92.15	13.93 15.47	N 110.90 P 101.44	13.19 15.05	N 108.78 P 95.63	13.17 15.39
	t=6.08	p=.000	t=4.96	p=.000	t=6.11	p=.000	t=7.53	p=.000
Bruininks Motor	N 59.79 P 55.13	11.02 11.33	N 59.97 P 49.59	10.47 14.28	N 60.82 P 56.70	10.25 11.99	N 60.60 P 53.05	10.17 13.02
	t=2.68	p=.008	t=4.78	p=.000	t=3.37	p=.001	t=5.08	p=.000

Table 17
Page 2

N = No Problem Group
P = Problem Group

Dependent	Language		Reading		Math		Physical Education	
Measure	Means	SD	Means	SD	Means	SD	Means	SD
Chronological	N 8.48	.84	N 8.49	.84	N 8.48	.84	N 8.50	.84
Age	P 8.83	.67	P 8.76	.64	P 8.71	.68	P 8.54	.57
	t=2.08	p=.038	t=1.45	p=.148	t=1.30	p=.194	t=.12	p=.905
Beery	N 9.50	3.07	N 9.47	3.10	N 9.57	3.06	N 9.43	3.07
Visual	P 7.84	2.70	P 7.85	2.18	P 7.08	2.28	P 6.17	1.17
Motor								
	t=2.61	p=.009	t=2.30	p=.022	t=4.06	p=.000	t=2.60	p=.010
WISC-R Full	N 116.03	15.55	N 116.03	16.15	N 116.28	15.55	N 115.35	16.22
Scale IQ	P 103.60	20.38	P 100.60	9.91	P 101.15	18.21	P 102.17	13.48
	t=3.75	p=.000	t=4.22	p=.000	t=4.70	p=.000	t=1.98	p=.049
Goodenough	N 98.84	16.41	N 98.92	16.39	N 99.32	16.31	N 98.68	16.57
Drawing	P 95.52	17.90	P 93.45	18.11	P 89.88	16.91	P 93.83	13.73
	t= .97	p=.334	t=1.44	p=.151	t=2.83	p=.005	t= .71	p=.477
Woodcock	N 106.15	13.14	N 106.37	13.01	N 106.34	13.22	N 105.46	13.43
Reading	P 94.24	13.64	P 87.85	8.82	P 92.42	10.23	P 95.17	15.82
	t=4.35	p=.000	t=6.27	p=.000	t=5.24	p=.000	t=1.85	p=.065
Woodcock	N 106.24	13.96	N 106.42	13.97	N 106.58	13.77	N 105.45	13.43
Math	P 95.80	17.11	P 90.35	14.02	P 92.08	16.13	P 95.17	15.82
	t=3.53	p=.000	t=4.99	p=.000	t=5.09	p=.000	t=1.85	p=.065
Woodcock	N 108.14	14.15	N 108.23	14.21	N 108.22	14.26	N 105.67	14.40
Language	93.56	15.15	P 88.50	9.08	P 93.19	13.02	P 94.17	14.70
	t=4.93	p=.000	t=6.13	p=.000	t=5.20	p=.000	t=1.94	p=.053
Bruininks	N 59.57	10.89	N 59.66	11.06	N 59.82	10.82	N 59.27	11.14
Motor	P 53.76	13.27	P 50.85	9.64	P 51.08	12.34	P 51.83	11.29
	t=2.52	p=.012	t=3.48	p=.001	t=3.91	p=.000	t=1.62	p=.106

c) Determination of whether the twenty seven MAP items can discriminate between the problem/no problem groups by conducting two tailed, t-tests for independent groups using each MAP item as the dependent measure for each of the eight problem categories. Table 19 provides a rank ordering (by t value from highest to lowest) of all items significant at the .01 level or less for each of the eight problem categories. T values and p values are noted in Table 19. Table 20 provides the corresponding means and standard deviations for each of the significant items in Table 19.

In the following discussion, the findings in each of the eight problem categories are presented. The discussion follows the following outline in each of the eight problem categories:

a) Do the groups represent different populations?
b) Results of analysis for MAP Total Score and Indices
c) Results of analysis for MAP items

1. Retained In School

a) Do the groups represent different populations?

As demonstrated in Table 17, In the Retained In School category, all of the dependent measures except chronological age demonstrate that the problem and no problem groups were distinct (p = .000 to .015).

b) Results of analysis of the MAP Total Score and Indices

Table 18 indicates that the means of the retained in school vs. not retained in school groups were significantly different on the MAP Total score, FI,VI,NVI,and CTI (p = .000 to .008). The scores on the two groups on the CI were not significantly different, but approached significance at p = .063.

c) Results of analysis of the MAP Item Scores

The items which differentiated between the retained in school vs. not retained in school groups were: figure ground, digit repetition, stamp, draw a person, supine flexion, block tap, stereognosis, and puzzles. These items include visual perceptual, verbal memory, fine motor, rapid movement patterns, motor strength, tactile, and sequential memory aspects of development.

2. Failed Teachers' Observations

a) Do the groups represent different populations?

In the Failed Teachers' Observations category, all of the dependent measures indicate that two distinct groups were being

Table 18
Two Tailed t-Tests For Problem vs. No Problem Categories Using MAP Total Score And MAP Indices As The Dependent Measure

*df=336

	Retention In School		Teacher Observations		Receiving Special Services		In A Special Class	
	N: not retained (n=291) P: retained (n= 47)		N: 50% or more "no problem" (n=311) P: 50% or more "problem" (n=27)		N: not receiving special services (n=201) P: receiving at least one service (n=137)		N: not in any special class (n=273) P: in a special class (n=65)	
Dependent Measure	Means	t values & p values	Means	t values & p values	Means	t values & p values	Means	t values & p values
MAP Total Score	N: 49.55 P: 45.11	t=4.08 p=.000	N: 49.35 P: 44.14	t=3.74 p=.000	N: 50.32 P: 46.91	t=4.47 p=.000	N: 49.76 P: 45.49	t=4.49 p=.000
Foundations Index	N: 49.04 P: 44.59	t=3.40 p=.001	N: 48.67 P: 45.57	t=1.83 p=.068	N: 49.18 P: 47.30	t=2.01 p=.046	N: 48.82 P: 46.72	t=1.81 p=.071
Coordination Index	N: 49.40 P: 46.72	t=1.86 p=.063	N: 49.34 P: 45.50	t=2.09 p=.037	N: 50.32 P: 47.13	t=3.17 p=.002	N: 49.99 P: 45.01	t=4.01 p=.000
Verbal Index	N: 50.18 P: 44.37	t=2.70 p=.007	N: 50.03 P: 41.75	t=3.02 p=.003	N: 51.93 P: 45.63	t=4.21 p=.000	N: 50.89 P: 42.98	t=4.24 p=.000
Non-Verbal Index	N: 49.85 P: 45.47	t=2.65 p=.008	N: 49.87 P: 46.39	t=1.46 p=.145	N: 49.87 P: 48.32	t=1.32 p=.187	N: 49.58 P: 47.83	t=1.19 p=.234
Complex Tasks Index	N: 49.21 P: 42.64	t=3.43 p=.001	N: 48.84 P: 42.02	t=2.77 p=.006	N: 50.27 P: 45.39	t=3.63 p=.000	N: 49.42 P: 45.49	t=4.49 p=.000

Table 18
Page 2

	Report Card Language		Report Card Reading		Report Card Math		Report Card Physical Education	
	N: no problem (n=313) P: failed language in any grade (n= 25)		N: no problem (n=312) P: failed reading in any grade (n=20)		N: no problem (n=312) P: failed math in any grade (n=26)		N: no problem (n=332) P: failed PE in any grade (n=6)	
Dependent Measure	Means	t values & p values	Means	t values & p values	Means	t values & p values	Means	t values & p values
Map Total Score	N: 49.42 P: 42.86	t=4.59 p=.000	N: 49.40 P: 41.52	t=5.00 p=.000	N: 49.28 P: 44.78	t=3.16 p=.002	N: 49.09 P: 40.27	t=3.06 p=.002
Foundations Index	N: 48.81 P: 43.55	t=3.03 p=.003	N: 48.84 P: 41.73	t=3.71 p=.000	N: 48.84 P: 43.40	t=3.19 p=.002	N: 48.57 P: 40.22	t=2.41 p=.016
Coordination Index	N: 49.56 P: 42.45	t=3.80 p=.000	N: 49.37 P: 43.61	t=2.75 p=.006	N: 49.18 P: 47.18	t=1.07 p=.284	N: 49.13 P: 43.43	t=1.51 p=.132
Verbal Index	N: 49.84 P: 43.56	t=2.20 p=.029	N: 50.09 P: 38.03	t=3.86 p=.000	N: 49.88 P: 43.29	t=2.35 p=.019	N: 49.65 P: 33.90	t=2.79 p=.006
Non-Verbal Index	N: 50.05 P: 39.11	t=5.15 p=.000	N: 49.77 P: 40.86	t=3.71 p=.000	N: 49.42 P: 47.07	t=1.09 p=.279	N: 49.36 P: 42.55	t=1.56 p=.119
Complex Tasks Index	N: 48.61 P: 44.28	t=1.69 p=.091	N: 48.71 P: 41.63	t=2.51 p=.013	N: 48.28 P: 44.78	t=3.16 p=.002	N: 48.36 P: 44.40	t= .78 p=.437

N: No Problem
P: Problem

Table 19

Items Rank Ordered By Ability To Discriminate Between All The Problem/No Problem Groups.
(Only Item Where p<.10 Or Less Are Included) Based On t Tests Between Problem/No Problem Groups n=338

RETAINED IN SCHOOL	TEACHER'S OBSERVATIONS	SPECIAL SERVICES	SPECIAL CLASS
Not Retained n=291 Retained n=47	Passed 50% n=311 Failed 50% n=27	Not Receiving n=201 Receiving n=137	Not Spec. Class n=273 In Special Class n=65
Figure Ground t=3.10 p=.002	Stereognosis t=3.38 p=.001	Articulation t=4.09 p=.000	Sentence Repet t=3.95 p=.000
Digit Repetition t=2.76 p=.006	General Info t=3.08 p=.002	Digit Repetition t=3.96 p=.000	Articulation t=3.60 p=.000
Stamp t=2.74 p=.007	Block Designs t=3.04 p=.003	Sentence Repet t=3.40 p=.001	Cage t=3.45 p=.001
Draw A Person t=2.59 p=.010	Cage t=2.89 p=.004	Romberg t=3.11 p=.002	General Info t=3.03 p=.003
Supine Flexion t=2.38 p=.048	Supine Flexion t=2.79 p=.006	Cage t=3.09 p=.002	Romberg t=2.98 p=.003
Block Tap t=2.31 p=.022	Puzzles t=2.48 p=.014	Block Tap t=2.67 p=.008	Block Tap t=2.56 p=.011
Stereognosis t=2.17 p=.051	Tongue Movements t=2.47 p=.014	Block Design t=2.56 p=.011	Draw A Person t=2.40 p=.017
Puzzles t=2.12 p=.035	Digit Repetition t=2.42 p=.016	Maze t=2.45 p=.015	Block Design t=2.39 p=.017
	Imitate Postures t=2.35 p=.020	Follow Direct t=2.10 p=.036	Digit Repetition t=2.32 p=.021
	Figure Ground t=2.35 p=.019	Hand Nose t=2.06 p=.041	Maze t=2.28 p=.023
	Romberg t=2.03 p=.043	Draw A Person t=1.92 p=.055	Tower t=2.17 p=.030
			Stamp t=2.02 p=.045

Table 19
Page2

FAILED REPORT CARDS

LANGUAGE	READING	MATH	PHYSICAL EDUCATION
Passed n=313 Failed n= 25	Passed n=318 Failed n= 20	Passed n=312 Failed n= 26	Passed n=332 Failed n= 6
Block Tap t=5.06 p=.000	Digit Repetition t=4.30 p=.000	Stereognosis t=3.17 p=.002	Kneel-Stand t=3.75 p=.000
Walks Line t=3.20 p=.002	Stamp t=3.66 p=.000	Cage t=3.00 p=.003	Follow Direct t=3.21 p=.001
Puzzles t=3.19 p=.002	Sentence Repet t=3.48 p=.001	Digit Repetition t=2.97 p=.003	Draw A Person t=2.35 p=.020
Cage t=3.03 p=.003	Figure Ground t=3.34 p=.001	Puzzles t=2.53 p=.012	Sentence Repet t=1.98 p=.048
Figure Ground t=2.95 p=.003	Kneel-Stand t=2.62 p=.009	Supine Flexion t=2.35 p=.020	Cage t=1.97 p=.049
Kneel-Stand t=2.27 p=.024	Romberg t=2.57 p=.011	Romberg t=2.26 p=.025	
General Info t=2.10 p=.036	Puzzles t=2.54 p=.012	Block Design t=2.21 p=.028	
Stereognosis t=2.03 p=.043	Block Tap t=2.52 p=.012	Kneel-Stand t=2.15 p=.032	
Articulation t=1.92 p=.055	Cage t=2.29 p=.023	Stamp t=2.05 p=.041	
	Articulation t=2.25 p=.025	Draw A Person t=2.02 p=.044	
	Stereognosis t=2.14 p=.033		
	Draw A Person t=2.11 p=.036		
	Hand Nose t=2.07 p=.039		

Table 20

Means And Standard Deviations For Item Scores On Problem/No Problem Groups From Table 20
Rank Ordered By Ability To Discriminate Between Groups

(Items, significant at .01 level or lower included)

	Retained In School			Teacher's Observations			Special Services			Special Class	
	Not Retained n=291	Retained n=47		Passed 50% n=311	Failed 50% n=27		Not receiving n=201	Receiving n=137		Not Spec. Class n=273	In Spec. Class n=65
	N	P		N	P		N	P		N	P
	Figure Ground			Stereognosis			Articulation			Sentence Repet	
M	52.20	42.77	M	50.65	39.24	M	53.79	44.04	M	51.34	39.95
SD	18.49	24.00	SD	15.99	24.48	SD	18.75	25.07	SD	19.83	25.05
	Digit Repetition			General Info			Digit Repetition			Articulation	
M	49.51	40.77	M	50.44	37.44	M	51.82	43.11	M	51.91	41.15
SD	19.60	23.09	SD	20.71	24.95	SD	17.23	23.23	SD	20.41	26.29
	Stamp			Block Designs			Sentence Repet			Cage	
M	51.77	43.96	M	47.93	33.74	M	52.36	44.44	M	51.11	42.98
SD	17.58	21.35	SD	22.79	28.08	SD	19.09	23.63	SD	15.77	21.80
	Draw A Person			Cage			Romberg			General Info	
M	49.36	41.47	M	50.34	40.39	M	51.03	44.16	M	51.10	42.28
SD	19.06	21.24	SD	16.33	25.10	SD	17.81	22.70	SD	20.33	24.00
	Supine Flexion			Supine Flexion			Cage			Romberg	
M	50.08	43.45	M	49.95	40.06	M	51.93	46.06	M	49.32	41.62
SD	17.25	20.45	SD	16.98	24.43	SD	14.95	19.93	SD	18.90	23.92
	Block Tap			Puzzles			Block Tap			Block Tap	
M	50.21	44.47	M	53.20	42.94	M	51.31	46.63	M	50.49	44.91
SD	14.98	20.38	SD	20.32	24.23	SD	13.84	18.27	SD	14.50	20.42
	Stereognosis			Tongue Movements			Block Design			Draw A Person	
M	50.54	44.76	M	53.43	44.09	M	49.47	42.87	M	49.50	43.06
SD	16.14	21.49	SD	18.08	26.06	SD	21.88	25.32	SD	18.53	22.73
	Puzzles			Digit Repetition			Maze			Block Design	
M	53.34	46.45	M	49.07	39.28	M	50.91	44.73	M	48.28	40.56
SD	20.24	23.35	SD	19.63	25.60	SD	21.74	24.22	SD	22.58	26.44
				Imitate Postures			Follow Directions			Digit Repetition	
			M	50.45	41.20	M	52.58	47.81	M	49.53	43.08
			SD	19.24	23.96	M	19.19	22.20	M	19.15	24.05
				Figure Ground			Hand-Nose			Maze	
			M	51.62	42.44	M	51.78	47.92	M	49.79	42.61
			SD	19.03	23.90	M	16.02	18.27	M	22.19	25.26
				Romberg			Draw A Person			Tower	
			M	48.90	40.70	M	49.94	45.80	M	49.43	42.82
			SD	19.73	24.02	M	18.31	21.04	SD	21.69	23.52
										Stamp	
									M	51.66	46.58
									SD	17.30	21.76

Table 20
Page 2

FAILED REPORT CARDS

	Language		Reading		Math		Physical Education	
	Passed	Failed	Passed	Failed	Passed	Failed	Passed	Failed
	n=313	n=25	n=318	n=20	n=312	n=26	n=332	n=6
	N	P	N	P	N	P	N	P
	Block Tap		Digit Repetition		Stereognosis		Kneel-Stand	
M	50.60	34.60	49.45	29.83	50.58	39.67	50.90	20.17
SD	14.46	24.45	19.30	26.73	16.16	23.73	19.88	20.26
	Walks Line		Stamp		Cage		Follow Directions	
M	43.87	28.06	51.58	36.40	50.36	39.87	51.12	24.33
SD	23.61	25.67	17.48	24.96	16.49	23.89	20.19	25.65
	Puzzles		Sentence Repet		Digit Repetition		Draw A Person	
M	53.39	39.76	50.15	33.28	49.23	37.06	48.59	29.83
SD	20.14	24.97	20.66	26.45	19.61	25.08	19.31	25.13
	Cage		Figure Ground		Puzzles		Sentence Repetition	
M	50.35	39.56	51.77	36.93	53.20	42.54	49.46	32.08
SD	16.47	24.27	18.93	24.59	20.28	24.65	21.15	28.63
	Figure Ground		Kneel-Stand		Supine Flexion		Cage	
M	51.77	39.88	51.08	38.95	49.81	41.33	49.80	35.75
SD	18.90	24.52	19.75	25.21	17.24	22.91	17.13	25.55
	Kneel-Stand		Romberg		Romberg			
M	51.06	41.54	48.95	37.08	48.96	39.71		
SD	19.74	24.82	19.67	25.26	19.72	23.98		
	General Info		Puzzles		Block Designs			
M	50.09	40.82	53.09	41.03	47.61	37.06		
SD	21.01	23.91	20.36	24.81	23.01	27.72		
	Stereognosis		Block Tap		Kneel-Stand			
M	50.27	43.10	49.96	40.78	51.04	42.19		
SD	16.44	22.94	15.13	24.40	19.77	24.54		
	Articulation		Cage		Stamp			
M	50.49	41.72	50.09	40.96	51.27	43.63		
SD	21.38	28.29	16.77	23.81	17.74	23.53		
			Articulation		Draw A Person			
M			50.51	39.15	48.88	40.87		
SD			21.34	29.71	19.03	24.08		
			Stereognosis					
M			50.23	41.85				
SD			16.49	23.58				
			Draw A Person					
M			48.82	39.38				
SD			19.12	24.15				
			Hand-Nose					
M			50.70	42.60				
SD			16.59	22.30				

N: No Problem
P: Problem

measured. All differences between groups were significant at p<.0001 level except chronological age which did appear to be different at this age at the .005 level.

b) Results of Analysis of the MAP Total Score and Indices

The mean scores of the passed teachers' observations vs. failed teachers' observations groups were significantly different on the MAP Total score, CI, VI, and CTI (p = .000 to .037). Their difference in performance on the FI approaches significance (p = .068) however the difference in the NVI performance was not significant.

c) Results of Analysis of the MAP Item Scores

The items which differentiated between the passed teacher observations and failed teacher observations were: stereognosis, general information, block designs, cage, supine flexion, puzzles, tongue movements, digit repetition, imitation of postures, figure ground, and romberg. These items represent the tactile, visual perceptual, oral motor, verbal memory, fine motor, motor strength, and balance domains of development.

3. Received Special Services

a) Do the Groups represent different populations?

As noted in Table 17, in the Received Special Services category, only age did not divide the two groups distinctly. All of the p values were less than .0001 except the Goodenough which was p = .004.

b) Results of analysis of the MAP Total Score and Indices

The mean scores of the receiving special services vs. not receiving special services groups were significantly different on the MAP total score, FI, CI, VI, and CTI (p = .000 to .046), as noted in Table 18. The difference in mean performance on the NVI was not significant.

c) Results of Analysis of the MAP Item Scores

The items which differentiated between the children who received special services, and those who did not were: articulation, digit repetition, sentence repetition, romberg, cage, block tap, block design, maze, follow directions, hand nose, and draw a person. It is interesting to note that the best three discriminators were language related items; it would be interesting to further analyze the data to see if a majority of the special services received were speech and language oriented. In addition several equilibrium items, fine motor items, motor planning items, and items requiring memory were discriminative.

4. In A Special Class

a) Do the Groups Represent Different Populations?

In the In A Special Class category all groups were distinct, as is clearly noted in Table 17. All categories were different at a highly significant level as indicated by p values less than .0001 in all categories except the Goodenough ($p = .007$) and Age ($p = .03$).

b) Results of Analysis of the MAP Total Score and Indices

The mean scores of the in a special class vs. not in any special class groups, as seen in Table 18, were significantly different on the MAP Total score, CI, VI, and CTI ($p = .000$). The difference in mean performance on the FI approached significance ($p = .071$) however mean performance on the NVI was not significantly different.

c) Results of Analysis of the MAP Item Scores

The items which differentiated between the children who were in a special class and those who were not as demonstrated in Tables 19 - 20, included: sentence repetition, articulation, cage, general information, romberg, block tap, draw a person, block design, digit repetition, maze, tower, and stamp. Several language items are included in this group, and it would be worth examining the special classes to see how many of them represent language delay classes. In addition to the language items, items measuring fine motor, balance, sequencing, visual perception, and motor planning are included.

5. Failed Report Card Language

a) Do the Groups represent different populations?

In the Failed Report Card Language category, only the Goodenough did not appear to be measuring distinct groups ($p = .334$). As demonstrated by Table 17, all the other groups were different a a highly significant or significant level, except age ($p = .04$).

b) Results of Analysis of the MAP Total Score and Indices

The pass vs. fail language groups mean scores were significantly different on the MAP Total score, FI, CI, VI, and NVI ($p = .000$ to $.029$). The difference in mean score performance on CTI approached significance (p = .091),as seen on Table 18.

c) Results of Analysis of the MAP Item Scores

The items which differentiated between the children who received average or above grades in language, from those who were below average, as demonstrated by Tables 18 - 19, included: block tap, walks line, puzzles, cage, figure ground, kneel-stand, general information, stereognosis, and articulation. Of note in this list is the lack of language items particularly in the top ranking positions. Rather, it appears from this study that non verbal cognitive items, and fine motor items discriminated between the problem/no problem groups in the report card language category better than the verbal items.

6. Failed Report Card Reading

a) Do the Groups represent different populations?

In the Failed Report Card Reading_category, age and the Goodenough did not appear to be measuring groups which were significantly different (p = .148, and .151), as seen in Table 17.

b) Results of Analysis of the MAP Total Score and Indices

The pass vs. fail report card reading groups mean scores were significantly different on the MAP Total score, FI,CI,VI, NVI, and CTI (p = .000 to .013), as noted in Table 18. Of the subtests, the Foundations Index, Verbal Index, and Non-Verbal Index were the best able to discriminate between the reading groups.

c) Results of Analysis of MAP Item Scores

The items which differentiated between the children who received grades below average from those who were average or above in reading were: digit repetition, stamp, sentence repetition, figure ground, kneel stand, romberg, puzzles, block tap, cage, articulation, stereognosis, draw a person, and hand nose. It is interesting to note that more individual items discriminated between the problem/no problem groups in this category than in the other categories, and the items cover a wide range of behavioral domains. The fact that both digit repetition, and sentence repetition were quite high on the list in interesting, and may indicate that an early ability to repeat verbal information is an antecedent to reading. It is also interesting to note the large number of high ranking neurological items which assess rapid alternating movement patterns, equilibrium, and tactile abilities. This may indicate that neurological maturity in the preschool years is a prerequisite for the development of adequate reading skills in the primary school years.

7. Failed Report Card Math

a) Do the Groups Represent Different Populations?

In the Failed Report Card Math category, in Table 17, it can be seen that all the groups were significantly different at the p = .0001 level, except the Goodenough in which the difference was still highly significant at .005, and chronological age (p = .194).

b) Results of analysis of the MAP Total Score and Indices

The pass vs. fail report card math groups mean scores were significantly different on the MAP Total score, FI, VI, and CTI (p = .002 to .019). These two groups were not significantly different on the CI or the NVI.

c) Results of analysis of the MAP Item Scores

The items which differentiated between the children who received grades in math that were below average from those that were above average were: stereognosis, cage, digit repetition, puzzles, supine flexion, romberg, block designs, kneel stand, stamp, and draw a person. The items on the MAP were not as good on the whole at predicting math as they were at predicting reading and language, as evidenced by lower t scores, and fewer items which discriminated at a significant level. It seems likely that, in part, an early precursor of math skills is neurological maturity since stereognosis, supine flexion, romberg, kneel stand, and stamp all discriminated at a significant level. Some of the other items which discriminated at a significant level were more expected, since face content appears more related to math, such as digit repetition, and block designs. It would be interesting to further investigate possible neurological antecedents of mathematical abilities.

8. Failed Report Card Physical Education

a) Do the Groups Represent Different Populations?

In the Failed Report Card Physical Education_category a number of dependent measures were not significantly different; however, it should be noted that the problem group is quite small (n = 6), and therefore all the data in this category is somewhat suspect.

b) Results of Analysis of the MAP Total Score and Indices

The pass vs. fail report card physical education groups had significantly different mean scores on the MAP Total score, FI and VI (p = .002 to .016). These two groups did not have significantly different means on the CI, NVI, or CTI.

c) Results of Analysis of the MAP Item Scores

There were five items which appeared to discriminate between the above average, and below average children in physical education, based on Tables 19 - 20. However, any conclusions regarding these two groups must be made with caution, since the number in the problem category was small (n = 6). The items which appeared to discriminate between the problem/no problem groups were: kneel-stand, follow directions, draw a person, sentence repetition, and cage. It was surprising that several of the gross motor items that were hypothesized to precede later physical education skills did not discriminate between the two groups: supine flexion, walks line, and stamp. However, two fine motor items, draw a person, and cage did discriminate. It is also interesting that two of the language items discriminated at a significant level.

9. Summary of Information Derived from T-tests

a) Do the groups represent different populations?

Table 17 indicates that the means of almost all of the problem vs. no problem groups were significantly different on all dependent measures related to aptitude (Beery, Full Scale IQ, Goodenough, Woodcock Reading, Woodcock Math, Woodcock Language, and Bruininks). Most of the groups were not however, different in chronological age. These results provide evidence that using these criteria to separate students into problem and no problem groups did represent a valid partitioning of subjects by low vs. normal aptitude. It is logical that in most cases age is not significantly different between the groups, since the problem/no problem groups were hypothesized to cover the entire age range.

b) Results of Analysis of the MAP Total Score and Indices

The t-tests provided evidence that the means of the MAP Total score (in all eight categories) and the means of most of the subtests were significantly different, and did discriminate between problem and no problem students up to four years after MAP administration. Although due to the large number of t-tests that were performed it is possible that some of the significant results could have occurred by chance, since so many of the tests are significant, it is unlikely that many were effected by this chance characteristic. In addition, the significance level was set at a low level, in most cases the probability of these findings occurring by chance is less than 1 in 1000. Use of t-tests in this instance was also less of a problem because the direction of the group differences on each variable was predicted in advance.

c) Results of analysis of the MAP Item Scores

The item which discriminated between the most dependent measures was Cage, which discriminated between seven of the eight problem/no problem categories. Digit repetition, and draw a person were also excellent items from this perspective, discriminating between six of the categories. Four items discriminated at a highly significant level between five of the categories: block tap, stereognosis, puzzles, and romberg. Six of the items discriminated significantly between four of the problem categories: block designs, figure ground, kneel-stand, stamp, sentence repetition, and articulation. Supine flexion and general information discriminated between three of the categories; and hand-nose, maze and follows directions discriminated between two of the categories. The seven remaining items only discriminated between one or none of the problem/no problem categories: tower, imitation of postures, tongue movements, walks line, finger localization, object memory, and stepping.

10. Supplementary Information Related to the Problem Categories

In addition to t-tests, two sets of correlations were run relating to the eight problem categories, reported in Tables 21 - 22. Table 21 reports the correlations between the MAP Total Score, and all of the eight problem/no problem categories. Table 22 compares the correlations between the MAP total score, the eight problem categories, and the standardized criterion measures.

It is interesting to note in Table 21 that most of the eight problem categories are not highly correlated. It is logical in fact that a child might have been retained in school, but is now doing well, or that because a child was put in a special class or was receiving special services, therefore his/her grades were average or above. Exceptions were high relationships between the category Needs Special Services and In A Special Class ($r = .48$) and relationships between report card grades in reading, language, and math ($r = .43$ to $.65$).

From the information in Table 22 it is quite interesting to note that the correlations between the criterion measures and the standardized measures only exceeded .40 one time (In a special class with the Woodcock Reading score). The largest correlations appeared consistently between Woodcock

Table 21
Intercorrelations of MAP Total Score, Eight Problem/No Problem Categories

	Retained In School	Teacher Observations	Needs Special Services	In A Special Class	Report Card Language	Report Card Reading	Report Card Math	Report Card P.E.	MAP Total Test
Retained In School	1.00	.13	.09	.15	.34	.41	.24	.14	.22
Teacher Observations		1.00	.25	.35	.21	.30	.28	.04	.20
Special Services			1.00	.48	.14	.20	.17	.12	.24
Special Class				1.00	.21	.29	.23	.05	.24
Report Card Language					1.00	.65	.43	.30	.24
Report Card Reading						1.00	.59	.25	.26
Report Card Math							1.00	.21	.17
Report Card P.E.								1.00	.16
MAP Total									1.00

Table 22
Intercorrelations of MAP Total Score, Eight Problem/No Problem Categories and, the Standardized Criterion Measures

	Beery Visual Motor	Full Scale IQ	Bruininks Motor	Goodenough Drawing	Woodcock Reading	Woodcock Math	Woodcock Language	Walker Total Teacher	MAP Total Score
Retained In School	.13	.25	.16	.17	.28	.31	.32	.07	.22
Teacher's Observations	.27	.27	.25	.20	.32	.26	.30	.22	.20
Needs Special Services	.29	.29	.18	.18	.30	.26	.32	.21	.24
In A Special Class	.25	.27	.27	.15	.41	.32	.38	.11	.24
Report Card Language	.14	.20	.14	.05	.23	.19	.26	.12	.24
Report Card Reading	.12	.22	.19	.08	.32	.26	.31	.21	.26
Report Card Math	.22	.25	.21	.15	.27	.27	.27	.13	.17
Report Card P.E.	.14	.11	.09	.04	.10	.11	.06	.06	.16
MAP Total Score	.21	.50	.39	.19	.36	.38	.35	.13	1.00

Language and Reading, and the problem categories. Of note, was the relatively large (by comparison to the other correlations) correlation of the MAP and WISC-R (.50), and the MAP and Woodcock measures (.35 to .38). The correlations between the WISC-R and the problem categories were much smaller than between the MAP and the WISC-R.

The preceding analyses answer Research Question #2, but the data analysis procedures used to answer research questions one and two did not take into account possible overlapping contributions of items, that is, that the same contribution might be provided by several of the items which correlated with each other. In the preceding analyses each item was looked at separately. Research Question #3 was asked next, to determine if intercorrelations between items would effect the set of items which contributed to the dependent measure scores.

Research Question #3.

What is the best set of MAP items to predict each of the following dependent variables: Retained in school, Failed teachers' observations, Needs special services, In a special class, and Failed report card grades?

In order to answer this question a step-wise multiple regression was run for each of the dependent measures first entering all 27 MAP items. This is referred to as Order A in the following discussion and Tables. In addition, a second step-wise multiple regression was run for the each of the dependent measures entering only the MAP items for each dependent measure related to item content. In this latter case, subjective determination of the relatedness of the item content to the dependent measure, and correlational values were used to make a determination of which items to enter. In the following discussion and Tables, Order B refers to the second regression where only items related in content were entered. Figure 2 shows the order in which the items were entered for the each of the regression analyses. Results of the step-wise regression analyses are demonstrated in Table 23.

Intercorrelations between the various items that were shown to be predictor items were calculated. Predictor intercorrelations have a very large role in determining which items best predict a given dependent variable. It is essential to try to identify relatively different and independent measures when the team of predictors is identified. The results of the intercorrelations for each set of predictor variables are noted in Tables 24 - 30. The following discussion includes references to correlations between items, where appropriate.

Table 23
Comparison of Results of Step-Wise Regression Analyses When Predictors Consider All 27 Items (A) vs. Based On Topic Content (B)

ORDER	Retained In School	Teachers Observations	Special Services	Special Class
ORDER A	Figure Ground R=.167	Stereognosis R=.18	Articulation R=.218	Sentence Repetition R=.210
	Digit Repetition R=.217	Cage R=.240	Cage R=.263	Cage R=.262
	Stamp R=.240	Tongue Movements R=.268	Digit Repetition R=.296	Articulation R=.292
	Supine Flexion R=.259	General Information R=.286	Romberg R=.316	Romberg R=.307
	Block Tap R=.270	Draw A Person R=.300	Follow Directions R=.328	Kneel-Stand R=.319
		Supine Flexion R=.313	Sequencing R=.341	Object Memory R=.330
		Stepping R=.325	Block Tap R=.351	Maze R=.342
			Maze R=.362	General Information R=.352
			Kneel-Stand R=.370	
ORDER B	Figure Ground R=.17	Stereognosis R=.18	Articulation R=.22	Sentence Repetition R=.21
	Digit Repetition R=.22	Cage R=.24	Cage R=.26	Cage R=.26
	Stamp R=.24	Tongue Movement R=.27	Digit Repetition R=.30	Articulation R=.29
	Supine Flexion R=.26	General Information R=.29	Romberg R=.32	Romberg R=.31
	Block Tap R=.27	Supine Flexion R=.30	Follow Directions R=.33	Maze R=.32
			Maze R=.34	

Table 23
Page 2

REPORT CARD STATUS

	Language	Reading	Math	Physical Education
ORDER A	Block Tap R=.263	Digit Repetition R=.228	Stereognosis R=.171	Kneel-Stand R=.201
	Puzzles R=.304	Figure Ground R=.282	Cage R=.235	Follow Directions R=.259
	Walks Line R=.339	Stamp R=.311	Sequencing R=.266	Maze R=.279
	Kneel-Stand R=.354	Sentence Repetition R=.326	Digit Repetition R=.294	Draw A Person R=.300
	Figure Ground R=.365	Cage R=.337	Kneel-Stand R=.310	Imitate Postures R=.311
	Cage R=.373	Kneel-Stand R=.345	Tongue Movements R=.323	Cage R=.320
	Draw A Person R=.385	Tongue Movements R=.353	Vertical Writing R=.334	Stamp R=.329
		Block Tap R=.361	Figure Ground R=.345	
			Maze R=.353	
ORDER B	Block Tap R=.26	Digit Repetition R=.23	Stereognosis R=.17	Not Analyzed Since n=6 in Problem Category
	Puzzles R=.30	Figure Ground R=.28	Cage R=.24	
	Walks Line R=.34	Stamp R=.31	Digit Repetition R=.26	
	Figure Ground R=.35	Sentence Repetition R=.33	Puzzles R=.28	
	Cage R=.36	Cage R=.34	Maze R=.29	

1. Retained In School

The team of predictors, each making a distinct contribution, that best predicted the first category, Retained in school were: figure ground, digit repetition, stamp, supine flexion, block tap. The best predictor was figure ground ($R = .167$), and when the effect of this variable was partialed out from the other variables,the item which added the most to predicting the variance in the dependent measure, Retained in school, was seen to be digit repetition. The multiple correlation for figure ground alone was .17, but when the pair of variables was considered the correlation was raised to .217. When all five predictor variables were entered, the correlation between the set of predictor variables and Retained in school was .27.

The squared multiple correlation corresponds to the proportion of the variance in the criterion score that is predictable from the set of predictor items. In this case of multiple correlation was .27.

It is interesting to note that the best predictor is a visual perceptual item, but one that has been found by some researchers to be highly correlated with intelligence measures. Digit repetition, also widely believed to be a general correlate of intelligence, does not have a high overlap with figure ground ($r = .02$), but contributed significantly to predicting the dependent measure. Stamp and supine flexion surprisingly entered into the equation next. These are gross motor items which are not widely known to be correlates of achievement in school; it is logical however, that these two variables would not overlap with the first two highly cognitive variables. The highly significant correlation between supine flexion and figure ground ($r = .17$), and between stamp and digit repetition ($r = .25$) was unexpected.

The simplest way to combine the information from the set of predictors is to express it as a simple linear composite in which each predictor score is given an appropriate weight, and the weighted scores are combined additively. The equation for Retained In School (X_1) is shown below:

$$X_1 = -.0022(\text{Figure Ground}) - .0018(\text{Digit Repetition}) - .0020(\text{Stamp}) - .0017(\text{Supine Flexion}) - .0017(\text{Block Tap}) + .61$$

Figure 2

Items which were Entered For Regression Analyses in Table 23

For Order A: All Twenty-Seven MAP Items Entered

Tower; sequencing; block designs; block tap; stereognosis; finger localization; object memory; puzzles; figure ground; draw a person; cage; vertical writing; hand-nose; romberg; stepping; walks line; supine flexion; kneel-stand; imitate postures; tongue movements; stamp; maze; general information; follows directions; articulation; sentence repetition; and digit repetition.

For Order B: Retained In School:

Figure ground; digit repetition; stamp; draw a person; supine flexion; block tap; stereognosis; puzzles.

For Order B: Failed Teachers' Observations:

Stereognosis; general information; block designs; cage; supine flexion; puzzles; tongue movements; digit repetition; imitation of postures; figure ground; romberg.

For Order B: Received Special Services:

Articulation; digit repetition; sentence repetition, romberg; cage; block tap; block design; maze; follow directions; hand-nose; draw a person.

For Order B: In A Special Class:

Sentence Repetition; articulation; cage; general information; romberg; block tap; draw a person; block design; digit repetition; maze; tower; stamp.

For Order B: Failed Report Cards Language:

General information; articulation; digit repetition; block tap; puzzles; figure ground; stereognosis; cage; sentence repetition; walks line; block designs; maze.

Figure 2 Page 2

For Order B: Failed Report Cards Reading:

Digit repetition; sentence repetition; figure ground; puzzles; block tap; articulation; block designs; stamp; romberg; cage; maze; stereognosis.

For Order B: Failed Report Cards Math:

Stereognosis; digit repetition; block designs; sentence repetition; cage; puzzles; romberg; block tap; maze; supine flexion; articulation.

Table 24

Intercorrelation Of Best MAP Predictor Items For Retained In School

	Figure Ground	Digit Repetition	Stamp	Supine Flexion	Block Tap
Figure Ground		.07	.09	.17	.16
Digit Repetition			.25	.07	.06
Stamp				.01	.07
Supine Flexion					.15
Block Tap					

Table 25

Intercorrelation Of Best MAP Predictor Items For Teacher Observations

	Stereog-nosis	Cage	Tongue Movements	General Information	DAP	Supine Flexion	Stepping
Stereognosis		.00	.03	.21	.13	.11	.14
Cage			.17	.04	.28	.14	.21
Tongue Movements				.10	.07	.10	.08
General Information					.04	.10	.05
Draw A Person						.14	.02
Supine Flexion							-.06
Stepping							

Table 26

Intercorrelations Of Best MAP Predictor Items For Needs Special Services

	Articulation	Cage	Digit Repetition	Romberg	Follow Directions	Sequencing	Block Tap	Maze	Kneel-Stand
Articulation		.21	.36	.12	.05	.19	.19	.15	.19
Cage			.07	.19	.02	.09	.14	.01	.04
Digit Repetition				.10	.06	.02	.06	.14	.14
Romberg					.06	.07	.07	.11	.10
Follow Directions						.10	.11	.04	.04
Sequencing							.21	.08	.12
Block Tap								.02	.04
Maze									.06
Kneel-Stand									

Table 27

Intercorrelations of Best MAP Predictor Items For In A Special Class

	Sentence Repetition	Cage	Articu- lation	Romberg	Kneel- Stand	Object Memory	Maze	General Inform.
Sentence Repetition		.14	.31	.14	.08	.06	.11	.39
Cage			.09	.19	.04	.02	.01	.13
Articulation				.12	.19	.08	.15	.10
Romberg					.10	.03	.03	.09
Kneel-Stand						.02	.06	.08
Object Memory							.12	.08
Maze								.10
General Information								

Table 28

Intercorrelations Of Best MAP Predictor Items For Report Card Language

	Block Tap	Puzzles	Walk Line	Kneel- Stand	Figure Ground	Cage	DAP
Block Tap		.07	.09	.04	.16	.26	.17
Puzzles			.00	.11	.24	.10	.08
Walks Line				-.03	-.04	.05	.02
Kneel-Stand					.09	.04	.14
Figure Ground						.06	.21
Cage							.28
Draw A Person							

Table 29

Intercorrelations Of Best MAP Predictor Items For Report Card Reading

	Digit Repetition	Figure Ground	Stamp	Sentence Repetition	Cage	Kneel-Stand	Tongue Movements	Block Tap
Digit Repetition		.07	.25	.29	.07	.14	.10	.06
Figure Ground			.09	.08	.06	.09	.06	.16
Stamp				.22	.05	.17	.07	.07
Sentence Repetition					.14	.08	.11	.11
Cage						.04	.05	.26
Kneel-Stand							.18	.04
Tongue Movements								.13
Block Tap								

Table 30

Intercorrelations Of Best MAP Predictor Items For Report Card Math

	Stereognosis	Cage	Sequencing	Digit Repetition	Kneel-Stand	Tongue Movements	Vertical Writing	Figure Ground	Maze
Stereognosis		.00	.01	.17	.09	.04	.08	.26	.17
Cage			.00	.07	.04	.05	.00	.06	.01
Sequencing				.02	.12	.02	.08	.25	.08
Digit Repetition					.14	.10	.02	.07	.13
Kneel-Stand						.18	.07	.09	.06
Tongue Movements							.03	.06	.11
Vertical Writing								.06	.03
Figure Ground									.13
Maze									

item which added the most to predicting the variance in the dependent measure, Retained in school, was seen to be digit repetition. The multiple correlation for figure ground alone was .17, but when the pair of variables was considered the correlation was raised to .217. When all five predictor variables were entered, the correlation between the set of predictor variables and Retained in school was .27.

The squared multiple correlation corresponds to the proportion of the variance in the criterion score that is predictable from the set of predictor items. In this case of multiple correlation was .27.

It is interesting to note that the best predictor is a visual perceptual item, but one that has been found by some researchers to be highly correlated with intelligence measures. Digit repetition, also widely believed to be a general correlate of intelligence, does not have a high overlap with figure ground ($r = .02$), but contributed significantly to predicting the dependent measure. Stamp and supine flexion surprisingly entered into the equation next. These are gross motor items which are not widely known to be correlates of achievement in school; it is logical however, that these two variables would not overlap with the first two highly cognitive variables. The highly significant correlation between supine flexion and figure ground ($r = .17$), and between stamp and digit repetition ($r = .25$) was unexpected.

The simplest way to combine the information from the set of predictors is to express it as a simple linear composite in which each predictor score is given an appropriate weight, and the weighted scores are combined additively. The equation for Retained In School (X_1) is shown below:

$$X_1 = -\ .0022(\text{Figure Ground}) - .0018(\text{Digit Repetition}) - .0020(\text{Stamp}) - .0017(\text{Supine Flexion}) - .0017(\text{Block Tap}) + .61$$

In other words, the multiple correlation (R) between the criterion score and the score that is obtained by combining the predictors with the specified weights above is .27. Thus 27% of the variance in the criterion measure can be predicted by this set of five predictor variables.

2. Retained In School

The step-wise multiple regression using Teachers' observations (as defined in research question #2 above) demonstrated in Table 22 that the best predictor was stereognosis ($r = .18$), and the following items also contributed significantly (had t values > 1.4) to the total R (.32): cage, tongue movements, general information, draw a person, supine flexion, and stepping. The correlation using seven predictor variables was significantly higher than with using only one.

It is interesting to note that the best predictor was stereognosis, a standard neurological item which assesses tactile integrative skills. Stereognosis was also the top ranking item on the list of items which discriminated significantly using t-tests as described above, in Research Question #2.

The linear equation for this criterion variable (X_2) was:

X_2 = - .0024(Stereognosis) - .0024(Cage) - .0015(Tongue Movements) - .0012(General Information) + .0014(Draw A Person) - .0015(Supine Flexion) - .0011(Stepping) + .51

3. Received Special Services

The multiple regression using Needs special services as the dependent variable had the following predictor variables: articulation, cage, digit repetition, romberg, follow directions, sequencing, block tap, maze and kneel-stand. The correlation ranged from .22 with one predictor to .37 with all nine predictor variables. All the predictor variables except sequencing and kneel-stand appear on the list of items which significantly discriminated between problem/no problem groups using t-tests (See Table 18).

It is interesting that Articulation is the top ranking item based on ability to discriminate between groups and is also the predictor variable which contributed the most to the predicted variance in Needs special services. Many researchers have felt that articulation is merely a motor skill, not well related to intelligence or general school functioning. This does not appear to be confirmed by these data.

The items in this group of predictor variables are quite varied representing all behavioral domains, speech, fine motor, verbal memory, equilibrium, receptive language, and motor planning.

The linear equation for this criterion, Needs Special Services (X_3), is:

X_3 = - .0033(Articulation) - .0029(Cage) - .0032(Digit Repetition) - .0027(Romberg) - .0022(Follow Directions) + .0032(Sequencing) - .0029(Block Tap) - .0019(Maze) +

.0020(Kneel-Stand) + 1.08

4. In A Special Class

The step-wise multiple regression using Needs a special class as the dependent variable had the following set of predictor variables: sentence repetition, cage, articulation, romberg, kneel-stand, object memory, maze and general information. The correlations ranged from .21 (with sentence repetition only) to .35 using the entire set of predictor variables.

This list corresponds closely to the list of items which discriminate between groups based on the t-tests. It is interesting to note that even though there is a high correlation between sentence repetition, and articulation ($r = .31$), apparently they contribute differently to the dependent variable since they both appear in the set of predictors. Noteworthy also is the high correlation between general information, and sentence repetition ($r = .39$). The predictor list includes items representative of all the developmental domains.

The linear equation which represents this criterion, In A Special Class (X_4) was:

X_4 = - .0019(Sentence Repetition) - .0029(Cage) - .0025(Articulation) - .0019(Romberg) + .0019(Kneel-Stand) + .0016(Object memory) - .0015(Maze) - .0017(General Information) +.64

5. Failed Report Card Language

The step-wise regression using Report card grades in Language as the dependent variable identified the following predictor variables: block tap, puzzles, walks line, kneel-stand, figure ground, cage and draw a person. The R values ranged from .26 if only one predictor is used, to .39 if the entire set of seven predictors is used. Interestingly, there is only one discrepancy between this list and the corresponding t-test table, the item draw a person, which is the last predictor variable.

It is interesting to note that Language did not have any of the MAP verbal items in the predictor variables. The best predictors, block tap and puzzles, were non verbal cognitive items. It is interesting to hypothesize that the early precursors of what later is measured in school as language performance may be cognitive non verbal tasks. Another hypothesis is that both language and non language cognitive tasks are predicted by the same cognitive tasks in the preschool years.

The linear equation for this criterion (X_5) was:

X_5 = - .0034(Block Tap) - .0014(Puzzles) - .0017(Walks Line) - .0014(Kneel-Stand) - .0015(Figure Ground) - .0016(Cage) + .0013(Draw A Person) + .55

6. Failed Report Card Reading

The regression analyses for Report card grades in Reading identify the following predictor variables: digit repetition, figure ground, stamp, sentence repetition, cage, kneel-stand, tongue movements, and block tap. The multiple R values are R = .23 using only the best predictor (digit repetition), to R = .36 using the set of eight predictors. With the exception of tongue movements which is the seventh of eight predictor variables, all of the predictor variables appear in the list of items which discriminate highly based upon the t-tests, although the order is slightly different.

It is interesting to note that the best predictor was a verbal memory item for reading grades but not for language grades. The important relative contribution of stamp (third in the list) was surprising, since it is a neurological item which is hypothesized to measure cerebellar functions, such as rapid alternating movement patterns.

The linear equation which expresses this criterion, Report Card Grades in Reading (X_6) was:

X_6 = .0018(Digit Repetition) - .0016(Figure Ground) - .0014(Stamp) - .0010(Sentence Repetition) - .0009(Cage) - .0010(Kneel-Stand) + .0010(Tongue Movements) - .0012(Block Tap) + .45

7. Failed Report Cards Math

The final regression analysis that was run was using the Report card grades in Math. The best set of predictor variables was seen to be: stereognosis, cage, sequencing, digit repetition, kneel-stand, tongue movements, vertical writing, figure ground, and maze. The multiple R is .17 using only the best predictor, stereognosis, but increased substantially to .35 using all nine predictor variables.

The overlap between the list of predictor variables and the list of items which discriminate between groups is not as close on this dependent measure. Items on the predictor list which do not appear in the t-test list are: sequencing, tongue movements, vertical writing, figure ground and cage.

The linear equation for Report Card Reading (X_7) was:

X_7 = - .0021(Stereognosis) - .0023(Cage) + .0024(Sequencing)
- .0017(Digit Repetition) - .0015(Kneel-Stand) +
.0012(Tongue Movements) + .0011(Vertical Writing)
-.0013(Figure Ground) + .0009(Maze) + .24

It is interesting that stereognosis was the best predictor for Report card grades Math, and for Teacher observations discussed above. This is the only item which was the best predictor variable for more than one dependent measure. The appearance of sequencing and digit repetition was expected in this list, however, stereognosis, cage, kneel-stand, and tongue movements were not hypothesized. Apparently there is a larger element of neuromotor functioning, than previously reported by research, which contributes to predicting this problem category.

When the all 27 MAP items were entered into the equation in a specified order, surprisingly, there were very few differences compared to the results when the MAP items related to content were entered: Order A being all 27 items, and Order B being related to content.

Some very slight changes were noted as follows: In Failed Teachers' Observations, Draw a person did not enter as early into the list; In Needs Special Services, Sequencing and Block tap exchanged positions with Maze; In A Special Class, Kneel-stand and Object memory were replaced by Maze; In Report Card Language, Kneel-stand was removed from the list; In Report Card Math, Sequencing and Kneel-stand were removed from the list and Maze was entered earlier. Report Card Physical Education was not rerun since the numbers in one group were so small (n = 6). None of the changes described above were felt to be highly meaningful.

In answering Research Question #3, it may be useful to look at the number of times each of the MAP items was indicated as a predictor variable. Cage was indicated the most, in six of seven dependent measures. Kneel-stand was indicated in five of the measures. The following items were indicated in four of the measures: Figure ground, Digit repetition, and Block tap. Tongue movement and Maze were indicated in three of the measures as predictor variables. Stamp, Supine flexion, Stereognosis, General information, Draw a person, Articulation, Romberg, Sequencing and Sentence repetition were indicated as predictor variables in two of the dependent measures. The following items were indicated only once: Stepping, Follows directions, Object memory, Puzzles, Walks line, and Vertical writing.

Research Question #4

What are the relative values of the MAP items; which of the MAP items can be considered the "best" or most useful items?

It would be an oversimplification to count the number of times each item appeared as a predictor variable, or ranked highly on a t-test, or correlated highly with a large number of criterion measures, in order to conclude which were the most "useful" items. In fact, the utility of the items is a many faceted issue; not only is it important how many of the criterion measures were correlated but which measures they were, and at what level the correlation was; not only is it important on how many of the dependent variables the item ranked significantly on the t-tests , but also at what level the discrimination between problem/no problem was significant; not only is it important to note how many times the item was a predictor variable, but also where it ranked in the list of predictor variables. Thus qualitative analyses are extremely important to conduct prior to making a definite determination of which items were the most useful. In addition, "usefulness" must be defined operationally.

However it is helpful for this study to look at the data provided in Figure 3. This figure provides a synthesis of the information supplied in Table 31. Table 31 notes the "value" of each item in terms of: the number of criterion measures with which it correlated, the number of WISC-R subtest with which it correlated at a significant level, the number of t-tests in which it ranked at a significant level, and the number of dependent measures for which it was a predictor in the regression analyses.

Figure 3 summarizes the information in Table 31 with regard to which are the "best" items. This figure has been put together based on the ranking of each of the items in each of the three categories, and provides only a gross overview of the value of the items. However, it appears that several of the items are "better" than the others in terms of predicting future behavior as measured by the criterion measures, and discriminating between problem/no problem groups. The better items include: cage, digit repetition, articulation, draw a person, block designs,stereognosis, block tap, puzzles, romberg, figure ground and maze.

One way this information could be used is in the development of a short form of the MAP. There certainly is a need for a short screening test to be used in well child clinics, and public school screenings. However the

Table 31

Summary of Items on Correlational Analysis, T-tests, and Regression Analysis

		Correlational Analysis		t-Tests	Regression
Index	Item Name	# of Significant Correlations	# of Significant WISC-R Subtests (of 10)	# of Significant t-tests (of 8)	# of Dependent Measures That Were Predictors (of 7)
Foundations Index	Stereognosis	14	10	5	2
	Finger Localization	4	3	1	0
	Vertical Writing	1	1	0	1
	Hand Nose	14	7	2	0
	Romberg	21	7	5	2
	Stepping	13	5	0	1
	Walks Line	5	1	1	1
	Supine Flexion	14	3	3	2
	Kneel-Stand	7	6	4	5
	Stamp	11	10	4	2
Coordination Index	Tower	13	9	1	0
	Cage	31	6	7	6
	Vertical Writing	1	1	0	1
	Walks Line	5	1	1	1
	Stamp	11	10	4	2
	Tongue Movements	5	3	1	3
	Articulation	26	10	4	2
Verbal Index	General Information	15	8	3	2
	Follow Directions	0	1	2	1
	Sentence Repetition	12	7	4	2
	Digit Repetition	28	10	6	4
Non Verbal	Sequencing	1	2	0	2
	Block Tap	21	9	5	4
	Object Memory	0	0	0	1
	Puzzles	22	10	5	1
	Figure Ground	10	10	4	4
Complex Tasks	Block Designs	26	10	4	0
	Draw A Person	14	7	6	2
	Imitate Postures	5	1	1	0
	Maze	22	8	2	3

Figure 3

Synthesis of Item Information Based Upon Correlative Analysis, T-tests, and Regression Analysis

Correlational Analysis	T-test Analysis	Regression Analysis
Cage	Cage	Cage
Digit Repetition	Digit Repetition	Kneel-Stand
Articulation	Draw A Person	Digit Repetition
Block Designs	Stereognosis	Block Tap
Puzzles	Romberg	Figure Ground
Maze	Block Tap	Maze
Romberg	Puzzles	Tongue Movements
Block Tap	Articulation	Stereognosis
General Information	Figure Ground	Romberg
Stereognosis	Block Designs	Sentence Repetition
Supine Flexion	Stamp	Supine Flexion
Hand Nose	Kneel-Stand	Stamp

development of such an instrument would require appropriate item analysis, pilot research, reliability and validity studies. Obviously this is beyond the scope of this dissertation; however, the item analysis completed to date provides a solid base for beginning development of a short screening test.

F. Research Question #5

How accurately do the recommended 5th percentile and 25th percentile cutpoints on the MAP predict a child's status as measured by the WISC-R, Woodcock Language, Woodcock Reading and Woodcock Math scores, and the problem categories defined in Research Question #2?

The purpose of this question was to further define the predictive value of the MAP by assessing individual children's status, rather than calculating what happens to the groups as a whole. As discussed in the Review of the Literature, classification data included: sensitivity, specificity, percent agreement, overreferral, underreferral, and referral rate, as well as raw frequency data regarding the numbers of children correctly and incorrectly identified.

In order to demonstrate how accurately the MAP predicted outcome status on the criteria, cross-tabulations of dichotomous MAP outcomes with the dichotomous outcome variables were performed. The results are expressed in 2 x 2 contingency tables. Two contingency tables were constructed for each of the twelve outcome criteria. The first used a cutoff score of 5th percentile on the MAP to determine risk status, the second used a cutoff score of 25th percentile. This allowed for analysis of each of the MAP's recommended cutpoints. A summary of the outcome of these analyses is reported in Tables 32 - 34. The following discussion is divided into subsections by each of the twelve criterion measures: the eight problem categories, and the four standardized measures. Following the report of the data is a section with conclusions relating to all of the twelve measures.

Table 32

Summary of Classificational Analyses of MAP Scores and Problem Categories
Using the 5th Percentile Cutoff Point On MAP

(referral rate is 4.4%)

	Sensitivity	Specificity	Percent Agreement	Overreferral Rate	Underreferral Rate
Retained In School	.11	.97	84.6	2.9	12.4
Teacher's Observations	.26	.97	91.7	2.4	5.9
Needs Special Services	.10	.99	62.7	<1	36.7
In a Special Class	.15	.98	82.2	1.5	16.3
Report Card Grades Language	.20	.97	91.1	2.9	5.9
Report Card Grades Reading	.25	.97	92.6	2.96	4.4
Report Card Grades Math	.12	.96	89.6	3.6	6.8
Report Card Grades Physical Education	.17	.96	94.4	4.14	1.5

Table 33

Summary of Classificational Analyses of MAP Scores and Eight Problem Categories Using the 25th Percentile Cutoff Point On MAP

(referral rate is 24%)

	Sensitivity	Specificity	Percent Agreement	Overreferral Rate	Underreferral Rate
Retained In School	.43	.79	73.96	18	7.9
Teacher's Observations	.52	.79	76.33	19.8	3.8
Needs Special Services	.32	.82	61.5	10.95	27.5
In a Special Class	.43	.81	73.37	15.68	10.95
Report Card Grades Language	.52	.78	76.33	20	3.6
Report Card Grades Reading	.60	.78	77.2	20.4	2.37
Report Card Grades Math	.50	.78	76.04	20	3.8
Report Card Grades Physical Education	.83	.77	77.2	22.5	<1

Table 34

Summary of Classificational Analyses of MAP with WISC-R and Woodcock Tests at the 5th Percentile and 25th Percentile Cut Points On MAP

	Sensitivity	Specificity	Percent Agreement	Overreferral Rate	Underreferral Rate
At 5th Percentile Cutoff					
WISC-R	.21	.97	89.6%	2.4%	7.99%
Woodcock Language	.23	.97	90.8%	2.4%	6.8%
Woodcock Reading	.20	.97	89.3%	2.4%	8.3%
Woodcock Math	.21	.97	89.94%	2.4%	7.6%
At 25th Percentile Cutoff					
WISC-R	.59	.80	77.8%	18%	4.1%
Woodcock Language	.53	.79	76.6%	19.2%	4.1%
Woodcock Reading	.54	.80	76.9%	18%	4.7%
Woodcock Math	.61	.80	78.1%	18%	3.8%

Interpretation of Classificational Analysis Proportions

A similar dilemma to that discussed in the section entitled "Interpretation of Predictive Correlation Coefficients" applys to the interpretation of the various classificational data including over and underreferral rates, sensitivity and specificity, and percent agreement. The point in the literature is consistently made that "some percentage of screening errors is unavoidable" (Lichtenstein 1984 p. 249), but the question is, what are acceptable percentages of hit rates in each category; what constitutes excellent vs. high vs. moderate predictive validity results?

In the absence of absolute numbers which correspond to specific descriptive modifiers, comparisons were made to similar studies involving young children who were screened in preschool or kindergarten, where the prediction interval was at least one year and results were presented in classificational form. A summary of these studies is shown in Table 35.

It must be cautioned that although this comparison is useful, it provides only a general sense of the value of particular results. To thoroughly evaluate these results factors providing possible sources of external invalidity must be carefully weighed.

What then, are considered "good" results for classificational analyses? What numbers can be used as guidelines for respectable predictive validity rates?

As Lichtenstein and Ireton point out (1984) after reviewing the results of a comprehensive survey of predictive validity studies:

> The results are humbling, indeed. Inspecting the validity data presented one finds that when sensitivity rates exceed .50, specificity rates are generally below .90 (meaning that over 10 percent of normal group children are referred) and often below .80 (over 20 percent of normals incorrectly referred). Furthermore, for most of these studies, prediction of high risk status proves to be substantiated by follow-up measure outcomes less than half of the time, i.e. efficiency of referral is generally below .50, it is rarely the case that sensitivity is also above .50, i.e. that more than 50 percent of target group children are identified." (p.253)

When the data from this study on the predictive validity of the MAP are compared to previous studies, it can be seen that

Table 35
Summary of Classificational Analyses Studies

Study	N	Prediction interval	Predictor	Criterion	Sensitivity	Specificity	Efficien of referr
Feshbach, Adelman, & Fuller (1974)	572	15 mos.	deHirsch Index	Gates Reading Test	.26	.93	.61
Feshbach, Adelman, & Fuller (1974)	585	15 mos.	Rating scale	Gates Reading Test	.30	.97	.83
Feshbach, Adelman, & Fuller (1977)	536	2 yrs.	deHirsch Index	Cooperative Reading Tests	.38	.91	.32
Feshbach, Adelman & Fuller (1977)	549	2 yrs.	Rating scale	Cooperative Reading Tests	.46	.93	.46
Feshbach, Adelman & Fuller (1977)	431	3 yrs.	deHirsch Index	Cooperative Reading Tests	.29	.91	.25
Feshbach, Adelman, & Fuller (1977)	451	3 yrs.	Rating scale	Cooperative Reading Tests	.43	.95	.49
Ireton & Thwing (1979)	287	1 yr.	Minnesota Preschool Inventory	Teacher Ratings	.60	.89	.30
Lichtenstein (1982)	428	1 1/2 yrs.	MPSI	Teacher Ratings	.63	.93	.62
Lichtenstein (1982)	428	1 1/2 yrs.	DIAL	Teacher Ratings	.54	.93	.59
Lichtenstein (1982)	296	2 yrs.	MPSI	Metropolitan Readiness Test	.56	.93	.62
Lichtenstein (1982)	296	2 yrs.	DIAL	Metropolitan Readiness Test	.46	.94	.61
Lindeman et al. (1967)	72	1-3 yrs.	Clinical interview	Teacher Ratings	.31	.91	.50
Lindquist (1982)	351	1 1/2 yrs.	DDST	Gates-MacGinitie Reading Test	.29	.89	.47
Satz, Friel, & Rudegair (1976)	151	2 1/2 yrs.	Abbreviated Satz Battery	Teacher rating, IOTA Word Recognition	.76	.71	.35
Stevenson, Parker, Wilkinson, Hegion & Fish (1976b)	152	2 1/2 yrs.	Teacher rating	WRAT, Stanford Achievement	.21	.93	.25
Wiske, Meisels, & Tivnan (1982)	78	1 1/2 yrs.	ESI	Academic grades	.92	.72	.40
Wiske, Meisels, & Tivnan (1982)	85	1 1/2 yrs.	ESI	Special Services	.81	.72	.41
Wiske, Meisels, & Tivnan (1982)	60	2 1/2 yrs.	ESI	Academic grades	1.00	.67	.32
Wiske, Meisels, & Tivnan (1982)	62	2 1/2 yrs.	ESI	Special Services	.81	.70	.48

Proportion of referred children falling in criterion measure problem group, i.e., likelihood that "refer" outcome is accura

From Lichtenstein & Ireton (1984) p. 250, 251 the MAP compares favorably to previous studies. MAP sensitivity and specificity figures are in most cases at least as high if not better than those described variously as respectable, good, excellent, high, etc. by the authors cited above. Thus in the following discussion, interpretations made regarding descriptive MAP data are seen to be consistent with those noted in similar studies.

1. Retained in School

It can be seen from Table 32 that when the 5th percentile cutpoint was used the level of sensitivity was .11, and the level of specificity was .97. At this cutpoint the MAP correctly identified 84.6% of the sample, and misclassified 15.4%. The referral rate was 4.4%. The overreferral rate of 2.9% was very low, and the underreferral of 12.4% was within acceptable limits.

When the 25th percentile cutpoint was used the levels of sensitivity and specificity were altered. The sensitivity was raised to .43, and the specificity level was lowered to .79. At this cutpoint, the MAP correctly identified 73.96% of the sample, with a referral rate of 24%. The overreferral rate was 18%, while the underreferral rate was 7.9%.

2. Failed Teachers' Observations

The classificational analysis of the teachers' observations using the 5th percentile as a cutpoint for the MAP demonstrated that the levels of sensitivity and specificity were .26 and .97 respectively. There was a high degree of overall agreement between the MAP and the teachers' observations, 91.7%. The overreferral and underreferral rates were quite low, at 2.4% and 5.9%. The referral rate was 4.4%.

When the cutoff point was raised to the 25th percentile, the sensitivity of the MAP increased significantly to .52, while the specificity decreased to .79, still well within acceptable limitations. The agreement between the MAP and teachers' observations accounted for 76.33% of the sample, while disagreement occurred on 23.67% of the sample. With an overreferral rate of 19.8%, and an underreferral rate of 3.8%, it can be seen that as the cutoff point is raised the percent of overreferrals increased, while the percent of underreferrals decreased slightly. The referral rate was 24%.

3. Received Special Services

The classificational analysis of the Received Special Services category, demonstrated in Tables 32 and 33, identified similar trends to those described above. In this category however, it must be remembered that a very high number of children actually received some extra services (40.5%). In this particular analysis all the children receiving any special service have been treated as a single group, and there has been no attempt to classify the service received with regard to the degree of "school problems" that the receipt of such services might indicate. Thus further analysis of this category is indicated.

When the 5th percentile cutpoint was used, the sensitivity level was .10, and the specificity was .99. Overall agreement on classification of children was 62.7%, with an extremely low overreferral rate of less than 1%, and a relatively high underreferral rate of 36.7%.

When the cutpoint was raised to the 25th percentile, the sensitivity was improved to .32, while the specificity decreased to .82. The agreement overall remained about the same at 61.5%. As expected, the overreferral rate increased to 10.95%, while the underreferral rate decreased significantly to 27.5%.

4. In a Special Class

In this sample, 65 of the 338 children were in special classes at the time of testing (19.23%). At the 5th percentile cutoff score on the MAP, the false positives were minimal (5) while the false negatives were high (55). Thus the sensitivity was seen to be .15, while the specificity was .98. Overall agreement between the MAP, and placement in a special class was 82.2%. The overreferral rate was quite low, at 1.5%, while the underreferral rate was moderate at 16.3%.

When the cutpoint was shifted to the 25th percentile in Table 33, the sensitivity increased substantially to .43, and the specificity remained quite good (.81). There was 73.37% agreement between the MAP and the criterion ratings. The referral rate was 24%, with an overreferral rate of 15.67%, and an underreferral rate of 10.95%.

With this criterion, it must be remembered that in many cases placement in a special class is subjective, and the definition of need may vary from school system to school system. Thus it would be helpful in future studies to further define this group of "special needs" children, to understand the composition of the problem group, and perhaps eliminate children who did not seem to fit the established definition of school problems.

5. Failed Report Card Language

Tables 32 and 33 report the findings for cutpoints of 5th percentile, and 25th percentile on the MAP for this criterion. Using the 5th percentile as the cutpoint, the sensitivity was seen to be .20, while the specificity was quite high at .97. The agreement rate was very high at 91.1%. Overreferrals were quite low at 2.9%, and underreferals were also low at 5.9%. There were twice as many false negatives (20), as false positives (10).

As would be expected, the sensitivity increased substantially when the cutpoint was raised to the 25th percentile, while the specificity dropped, though remaining within acceptable standards (.52, and .78 respectively). The overreferral rate increased substantially to 20%, while the underreferral rate decreased somewhat to 3.6%. The overall referral rate is 24%, with 76.33% agreement between the MAP and the criterion.

6. Failed Report Card Reading

The results of the classificational analysis of the MAP and the report card reading grades are reported in Tables 32 and 33. Reading had fewer false negatives (15 at the 5th percentile cutoff, and 8 at the 25th percentile cutoff), than any of the other criterion measures except report card grades in physical education, where the problem group is small (n = 6). From this perspective reading appears to be the criterion on which the MAP is least likely to "miss" children.

Further analysis of the data using the 5th percentile as a cutpoint demonstrated that the sensitivity of the MAP was .25, while the specificity was quite high at .97; the overreferral rate was very low at 2.9%, as was the underreferral rate of 4.4%. The agreement between the MAP and the criterion was high (92.6%).

When the cutpoint was raised to the 25th percentile, there were seven fewer false negatives (approximately 50% less, a substantial difference). The sensitivity increases to .60, while the specificity is .78, both within acceptable limitations. The overreferral rate increased to 20.4%, and the underreferral rate decreased to 2.37%. The overall referral rate was 24%, with 77.2% of the sample in agreement between the MAP rating and the criterion rating.

7. Failed Report Card Math

The analyses for this criterion are reported in Tables 32 and 33. At the 5th percentile cutpoint, sensitivity of the MAP was .12, and specificity was .96. The total referral rate was 4.4% with 3.6% overreferrals, and 6.8% underreferrals. The rate of agreement between the MAP and math grades was 89.6%.

As expected, when the cutpoint was raised to the 25th percentile, the actual numbers of false negatives decreased substantially from 23 to 13, however the overreferral rate increased to 20%, reflecting an increase in false positives. The sensitivity at this cutpoint was .50, and the specificity was .78. The agreement rate was 76.04% considering the total referral rate of 24%.

8. Failed Report Card Physical Education

Any analysis made with the groups in this category must be cautious since the number in the problem category is only six. The reader is warned to use caution in the interpretation of results.

When the 5th percentile was used as a cutpoint, the sensitivity was .17, while the specificity was .96. There was 94.4% agreement between the MAP and the Physical Education grades, with a referral rate of 4.4%. The overreferral rate was 4.14%, while the underreferral rate was 1.5%.

When the cutpoint is raised to the 25th percentile, the sensitivity increased to .83 (only one child is misclassified as a false negative), and the specificity decreased to .77, both figures being well within acceptable limitations. The overreferral rate then was 22.5%, while the underreferral rate was less than one percent.

The classificational analysis studies that were completed with the WISC-R and Woodcock standardized measures are and summarized in Table 34. As might be expected due to the standardized nature of these criterion measures, the levels of sensitivity and specificity are somewhat higher than they are with the more subjective criterion measures described above. With the following analyses, the cutpoint between problem and no problem on the standardized measures was set at the score which differentiated the bottom 10% of the sample. Specifically the cutpoints were:

WISC-R: cutpoint = Full scale score of 95 (bottom 10% of sample)

Woodcock Language: cutpoint = Derived Standard Score of 87 (bottom 10% of sample)

Woodcock Reading: cutpoint = Derived Standard Score of 88 (bottom 10% of sample)

Woodcock Math: cutpoint = Derived Standard Score of 85
(bottom 10% of sample)

9. WISC-R

On the WISC-R utilizing a cutpoint of 5th percentile on the MAP, a sensitivity of .21, and a specificity of .97 were noted. At this cutting level, the overall agreement between the MAP and the WISC-R was 89.6% with an overreferral rate of 2.4%, and an underreferral rate of 7.99%.

When the cutpoint was raised to the 25th percentile, the sensitivity was raised significantly to .59, while the specificity was maintained at a very good level (.80). The overall percent agreement was seen to be 77.8%. The overreferral rate was raised to 18%, while the underreferral rate dropped slightly to 4.1% Interestingly, the WISC-R appears to have the highest levels of sensitivity and specificity of all the dependent measures that were examined in this study.

10. Woodcock Language

The classificational analysis of the Woodcock Language Scores using the 5th percentile as the cutpoint on the MAP demonstrated a sensitivity of .23, and a specificity of .97. The overall agreement between the MAP and the Woodcock Language was 90,8% with an overreferral rate of 2.4%, and an underreferral rate of 6.8%.

When the cutpoint of the MAP was raised to the 25th percentile, the sensitivity was increased to .53, while the specificity remained good at .79. The overall agreement rate dropped somewhat to 76.6%, and there was an overreferral rate of 19.2%, but an underreferral of only 4.1%.

11. Woodcock Reading

Table 33 summarizes the results of classificational analysis of the MAP and the Woodcock Reading scores. Using the 5th percentile as a cutpoint, the sensitivity of the MAP was seen to be .20, and the specificity was .97. This resulted in an overall agreement rate of 89.3%. The overreferral rate was 2.4%, while the underreferral rate was 8.3%.

When the cutpoint on the MAP was raised to the 25th percentile, the sensitivity was raised to .54, and the specificity remained within acceptable limits at .80. The overall percent agreement was lowered to 76.9% which reflected an overreferral rate of 18%, and an underreferral rate of 4.7%.

12. Woodcock Math

The classification analysis of the MAP and the Woodcock Math scores is also summarized in Table 33. As demonstrated in this table, the sensitivity and specificity of the MAP at the 5th percentile cutpoint was .21 and .97 respectively. The overall percent agreement was 89.94%, while the overreferral rate was 2.4% and the underreferral rate was 7.6%.

When the MAP cutpoint was raised to the 25th percentile, the sensitivity increased significantly to .61, while the specificity remained at a high level (.80). The percent agreement was 78.1% with an overreferral rate of 18%, and an underreferral rate of 3.8%.

13. Summary of Findings of Classificational Analysis

It can be observed in Table 32 that when the cutoff point on the MAP is placed at the 5th percentile, the sensitivity of the eight problem categories ranged from .10 to .26. Specificity ranged from .96 to .99. Percent agreement ranged from 62.7 to 94.4. The overreferral rate ranged from less than 1 % to 4.14 % while the underreferral rate ranged from 1.5% to 36.7%.

When the cutpoint of the MAP was changed to the 25th percentile corresponding changes were seen as expected in the sensitivity, specificity, percent agreement, overreferral rate and underreferral rate of the eight problem categories. Sensitivity ranged from .32 to .83. Specificity ranged from .77 to .82. Percent agreement ranged from 61.5% to 77.2%. Overreferrals were at a rate of 10.95 to 22.5%, while underreferrals ranged from less than one percent to 27.5%.

It should be noted in particular with the false positive classification that the figures are inflated somewhat since the eight dependent measures are interrelated. For example, in reviewing the data (noted in Table 36) it was found that of the 61 total false positives in Retained in school (at 25th percentile cutpoint), 34 were found to be true positives on one of the other 7 dependent measures. The number of false positives that were found to be true positives, by looking at their classification on the other 7 dependent variables ranged from a low of 10 (in Received Special Services) to a high of 49 (in Report card Physcial Education. The number of true positives was 27 across all categories. Thus for the child who was not retained in school but did poorly on the MAP, 34 were found to either be In a special class, Failing on report cards, etc.

The classificational analysis data for the four standardized measures are summarized in Table 34. At the 5th

Table 36

Analysis of False Positives in Eight Problem/No Problem Categories

	Retained In School	Teacher' Observations	Needs Special Services	In A Special Class	Report Card Language	Report Card Reading	Report Card Math	
Apparent # of False Positives in Category	61	67	37	63	68	69	68	76
# That Appear in One of the Other 7 Categories as a Problem	34	40	10	26	41	42	41	49
# True False Positives in Each Category	27	27	27	27	27	27	27	27

percentile cutpoint the sensitivity ranged from .20 to .23, while the specificity was .97 for all four measures. The percent agreement was from 89.3% to 90.8%. The overreferral rate was the same for all measures: 2.4%, and the underreferral rate ranged from 6.8% to 8.3%.

At the 25th percentile cutpoint, the sensitivity increased significantly for all measures to range from .53 to .61, while specificity remained high for all measures, .79 to .80. The overall agreement rate was 76.6% to 78.1%. The overreferral rate was 18% to 19.2%, and the underreferral rate was 3,8% to 4.1%.

It is noteworthy that there appears to be consistency between these four measures. In addition, the values are somewhat higher than reported above for the eight other criterion measure, which are more subjective in nature.

It is essential when interpreting the classification data from this study,, to remember that the criterion measures to which the MAP was intended to predict are not perfect measures of school success or failure. In particular, it is important to remember that the first eight criterion measures described in this part of the study are all subjective, dependent upon judgments from teachers or school records. The lack of standardized criteria may certainly have had a large influence on the outcome of the study. The ability of the MAP to accurately predict standardized test scores was better than its ability to predict the non-standardized criteria.

The results of this study reflect results obtained with the MAP when a specific decision rule (e.g. 5th percentile and 25th percentile cutoff points) was utilized. Use of independently derived decision rules in which cutoff points were not devised after the fact (so as to maximize the hit rates obtained) consistently provides data that makes the test examined appear more valid. If further studies were completed choosing specific cutpoints, it is likely that the sensitivity and specificity levels could be increased.

In addition, the predictions examined in this study were made four years before the criterion measures were administered. If the amount of time were decreased, it is likely that all the relevant data would be improved.

In addition, it must be reemphasized that there is no magic solution to what is known as the "prediction predicament" (Lichtenstein and Ireton, 1984). It is possible to alter the decision criteria so as to eliminate either false positives, or false negatives in most screening tests; however, they can not both be eliminated at the same time. As one rate increases, the other decreases. A final decision regarding what cutpoint a particular system wishes to utilize must revolve around the screening philosophy of the agency involved. It is likely that a large number of false positives will have to be screened in order to "flag" all the potential children with problems, thus increasing the total numbers screened; on the other hand, if the goal is to reduce the total number of children screened, i.e. to reduce initial costs, it is likely that numerous false negative errors will result. A further decision that needs to be made by the screening agency is whether false positives, or false negatives "cost" more.

However, keeping the above guidelines in mind, it can be concluded that the preidication rates for the MAP compare quite favorabley to those of similar studies. Sensitivity, specificity and percent agreement rates are well within acceptable standards (Lichtenstein, 1984; Lemerand, 1985). In general both the overreferral and underreferral rates are excellent.

If the scoring system of the MAP were ever revised it would be important to determine from this predictive validity information if the sensitivity rates could be enhanced through item weighting, and other statistical techniques. This could provide a means for reducing the numbers of false negatives.

Limitations of this Study

The limitations of this study are noted below. These limitations cover the following areas: 1) History; 2) Instrumentation; 3) Mortality; and 4) Sample Selection. The following areas which were found to be limitations in many of the studies reviewed were not problems in this study: practice effect, time, criterion contamination, multiple measure interference, and replicability.

1. History

In any study which is longitudinal, there will be numerous external factors which impose limitations on prediction (Coons et al., 1982; van Doornick, 1978; Keogh, 1977; Lewis, 1980; Meisels, 1984;). For example, the evaluation of the outcome of the study may be confounded by a treatment effect. The predictions from the original MAP scores to the criterion measures may have been affected by support services some children received during the time interval between pre-testing and post-testing. Another example is that children who were positive on one of the criterion measurements, such as retained in school, might not then show up as a positive in another category such as failed report card grades.

Although an attempt was made to be as comprehensive as possible in denoting possible confounding effects of history through parent forms, teacher forms, and independent review by Field Staff, it is likely that numerous confounding effects of history occurred which were unaccounted for.

2. Instrumentation

An attempt was made in the selection of criterion measures to choose measures that were standardized as well as possible with good reliability and validity evidence. However, there are never perfect measures of any criteria. Particularly the use of non-standardized measures represents a possible source of bias. The outcome of the children on these measures is dependent on one person's perceptions, or one school systems philosophy of retention in school, or need for specialized service. Thus the ability of the MAP to predict outcome of children four years later in this study was effected by the fact that there is no perfect measure of school success/failure in primary school years.

3. Mortality

It is expected that with any longitudinal study serious confounding effects may occur as a result of "loosing" children. In this study the demographic composition of the sample seems quite similar to the original standardization group on most variables. However, the original standardization was done with 1204 children; follow up permission was obtained for 800 at the time of testing; 400 children were unable to be tracked over the entire four year period between pre-test and post-test. Of the 400 remaining children, 338 were retested for this study (the rest having moved to a geographic area where they could not be reached). This does represent a serious mortality rate; although demographic variables do not seem to be different, it is quite possible that the children who were not tested were different from the children who were tested in some significant ways that were not documented statistically. This represents the most serious potential limitation of this study.

4. Sample Selection

In the original sample of the MAP standardization project, geographic region, race, community size, age, and sex were representative of the U.S. Census Bureau statistics. However, there was a skew in the sample towards the higher income/ higher professional categories/ more educated socio-economic variables. This skewed sample was also reflected in the sample of this study. It is possible that this skewed characteristic might have effected the outcome of this study, although the direction of the effect can not be predicted.

Recommendations for Further Study

There are two types of further study that are recommended. The first would be corrections of some of the limitations of this study. The second would be extentions of the MAP materials based on these intial promising validity results.

There are three studies that would be recommended to correct the limitations of the study described in this report.

1. Comparative Analysis: It would be useful to compare the longitudinal validity of the MAP to the other most commonly used preschool tests, the DDST, the CIP and the DIAL-R. Preferably this would be done with the same sample of children followed over two to three years. Then the comparative usefulness of the MAP to the other preschool screening tools could be established.

2. Analysis of "Problem" children: Although a number of children who were thought to be at risk for school related problems were followed through this study, it would be quite helpful to study a large group of children with known dysfunction in the moderate range. This would further validate the use of the MAP with this population of children.

3. Normal Distribution: It would be helpful to replicate this study with a sample of children representing a normal distribution on socio-economic variables. As mentioned previously this sample was skewed in the direction of higher income, and higher educational levels. Replication with a sample more representative of the U.S. population on SES variables is indicated.

In the second area, extensions of the MAP, several types of research would be useful:

1) Development of the Mini-MAP: The development of a shorter version of the MAP, which would be cost effective, and would be designed to overidentify children is indicated. That way in a much shorter time, 30-35% of the children who may be at risk could be flagged, and only those children who were flagged would then receive the longer, and more complete MAP. This Mini-MAP would be intended to have no false negatives, and would be compiled based upon analysis of predictive information from the MAP.

2) Development of the Diagnostic MAPS: Once children have been identified as potentially at risk for problems, it would be helpful to have well standardized and reliable diagnostic tools with which to examine them. There are a paucity of assessments for this age group. Based upon the predictive information from this study, and analysis of predictive information on the research edition of the MAP (100 children who were examined on the Research Edition of the MAP were also post-tested on the criterion measures described in this study, but were not included in the analyses described in this report) several Diagnostic MAPS could be constructed, a neuro-motor test, a verbal test, and a non-verbal cognitive assessment for children aged two to six.

3) Development of the MAP Training Guide: Once the children have been identified and assessed in depth, remedial programs aimed at alleviating their difficulties should be undertaken. At the moment most professionals rely on their past experience to develop these programs. It would be quite helpful to the field to develop an in-depth Training Guide which could be used in treatment planning.

Conclusion

The present study utilized numerous methods of statistical analysis to approach the question of predictive validity. The results are given extra credibility since four different methods of analyzing data were used, and all essentially demonstrated the same positive outcome. The results of this study indicate that the predictive validity of the Miller Assessment for Preschoolers compares favorably to other similar predictive validity studies. It can consistently and correctly identify the majority of children who later have difficulty in primary school, with relatively few misclassifications.

Hopefully, the establishment of clear predictive validity will foster the widespead use of the MAP in preschool screening programs and thereby facilitate the detection and possible remediation of potential school problems.

References

Borg, W.R., & Gall, M.D. (1983). Educatinal Research: An Introduction (4th ed.). New York: Longman.

Coons, C., Frankenburg, W., Gay, E., Fandal, A., Lefly, D., & Ker, C. (1982). Preliminary results of a combined developmental/environmental screening project. In N. Anastasiow, A. Fandal, & W. Frankenburg (eds.), Identifying the developmentally delayed child. Baltimore: University Park Press.

Fox, D.J. (1969). The Research Process in Education. New York: Holt, Rinehart and Winston, Inc.

Keough, B. (1977). Early ID: Selective perception or perceptive selection? Academic Theory, 12, 267-274.

Lewis, A. (1980) The early identification of children with learning difficulties. Journal of Learning Disabilities, 13, 51-57.

Lichtenstein, R., & Ireton, J. (1984). Preschool screening: Identifying young children with developmental and educational problems. New York: Grune & Stratton, Inc.

Meisels, S.J., Wiske, M.S., & Tivnan, T. (1984). Predicting school performance with the Early Screening Inventory. Psychology in the Schools, 21, 25-33.

Sax, G. (1968). Empirical Foundations of Educational Research. Englewood Cliffs, N.J.:

van Doornick, W,. (1978). Prediction of school performance from infant and preschool developmental screening. In W. Frankenburg (Ed.), Proceedings of the Second International Conference on Developmental Screening. Denver: JFK Child Development Center.